SAVOIR & FAIRE EN PRÉPAS

COLLECTION DIRIGÉE PAR KARINE BEAURPÈRE

MATHS

MP/MP*

Thierry **Legay**

- Les **savoirs essentiels** du programme
- Les **problématiques les plus classiques**
- Les **méthodes** de résolution les plus utiles
- Des **exercices d'entraînement** minutieusement choisis
- Des **corrigés détaillés** de tous les exercices

ellipses

SAVOIR & FAIRE
EN PRÉPAS

COLLECTION DIRIGÉE PAR KARINE BEAURPÈRE

MATHS

MP/MP*

Thierry **Legay**

Professeur de chaire supérieure
au lycée d'Arsonval de Saint-Maur-des-Fossés

Ancien élève de l'École normale supérieure de Saint-Cloud

ellipses

Présentation de la collection

La collection « Savoir et Faire en Prépas » est destinée aux étudiants des Classes Préparatoires aux Grandes Écoles.

Chaque ouvrage de la collection s'adresse à un niveau donné d'une filière précise, pour une matière spécifique.

Cette collection est conçue dans le but de **pouvoir parfaitement utiliser son cours**. C'est pourquoi cette collection met en avant les savoirs théoriques essentiels mais elle met surtout en lumière la façon dont sont utilisés ces savoirs au travers de nombreuses méthodes.

Cette collection démontre finalement que **la résolution d'une grande quantité de problèmes se fait avec assez peu de savoirs théoriques mais beaucoup de savoir-faire**.

Aussi chaque ouvrage de la collection se divise en trois grandes parties.

• La première partie liste les **savoirs essentiels du cours**. Une fois ces savoirs exposés, une ou des remarques peuvent venir compléter ce qui a été exposé. J'insiste sur le fait qu'un savoir sera toujours utilisé au moins une fois dans la partie « Savoir-faire ».

• La seconde partie met en avant les **savoir-faire à connaître impérativement**, tout ceci s'organisant autour de problématiques bien identifiées et illustrées avec de nombreux exemples. Cette partie se termine systématiquement par des exercices d'entraînement qu'il est important de chercher activement afin que la lecture du corrigé ne soit pas inutile.

• La troisième partie est consacrée aux **corrigés des exercices**. Ces corrigés sont suffisamment détaillés pour être bien compris de tous les étudiants.

Finalement, travailler cet ouvrage est la garantie d'un apprentissage efficace du cours afin d'être le plus performant possible le jour du concours.

Karine Beaurpère

Comment utiliser ce livre ?

Dans la **partie « Savoirs »**, tous les savoirs sont identifiés à l'aide d'un codage et d'une appellation. Ce codage est le suivant : [S + numéro du thème + numéro du savoir dans ce thème].

Par exemple, **[S8.3]** représente le savoir n° 3 du thème 8.

C'est ce codage qui sera utilisé dans la partie « Savoir-faire » pour faire référence à un savoir précis.

Dans la **partie « Savoir-faire »**, on retrouve les mêmes thèmes que ceux de la partie « Savoirs », dans le même ordre.

Tout y est organisé autour de problématiques.

Une fois une ou plusieurs méthodes exposées rapidement pour une même problématique, on rentre dans le détail des différents savoir-faire liés à la problématique étudiée.

Comme ci-dessus, les savoir-faire sont identifiés à l'aide d'un codage et d'une action décrivant ce savoir-faire.

Par exemple, **SF3.5** représente le savoir-faire n° 5 du thème 3.

Chaque savoir-faire est ensuite illustré d'un ou plusieurs exemples, toujours étudié(s) en détail et parfois complété(s) de différentes remarques.

La lecture de la table des matières, située en fin de livre, peut être précieuse car elle permet, d'un seul coup d'œil, de visualiser, par thème, toutes les problématiques et les savoir-faire qui y sont reliés.

Enfin, une fois toutes les problématiques exposées, c'est à vous de jouer !

Pour cela, plusieurs exercices minutieusement choisis sont proposés.

Le codage des exercices relève du même principe que ce qui a été vu plus haut. De plus, en haut à droite de l'énoncé sont précisés les savoir-faire utilisés dans la résolution de l'exercice.

Il est absolument essentiel de chercher activement ces exercices afin que la lecture du corrigé ne soit pas inutile.

L'idéal est d'utiliser cet ouvrage régulièrement dans l'année, en particulier à la fin de chaque chapitre, afin de faire le point sur les savoirs et les savoir-faire.

L'étude de la première partie permet en effet de voir si l'**on a appris et retenu les « bons savoirs »**.

Puis, l'étude de la seconde partie permet de voir si l'**on a bien cerné les différentes problématiques liées à un thème donné** et surtout, cette étude permet de voir si l'**on a bien en tête les savoir-faire répondant à une problématique donnée**.

Le travail sur les exercices proposés doit finalement permettre de vérifier que les savoirs et les savoir-faire sont bien acquis.

Quelques conseils pour bien apprendre

Il est important d'**apprendre régulièrement** son cours et d'**être actif** lors de cet apprentissage.

Cela implique donc de reprendre ses notes très rapidement après le cours fait au lycée, pour ne pas accumuler de retard dans les apprentissages. Il est par exemple essentiel d'arriver à un cours en étant bien évidemment au point sur le (ou les) cours précédent(s).

Mais cela signifie aussi une vraie implication dans cet apprentissage. Pour cela, il faut se poser un certain nombre de questions afin de ne pas subir son cours. Voici quelques questions que vous pouvez vous poser régulièrement, notamment en relisant votre cours mais aussi en travaillant les corrigés des exercices et des devoirs.

– Ai-je bien tout compris à cette définition, cette propriété ou ce théorème ?
– Cette définition est-elle essentielle en pratique ?
– Cette propriété est-elle vraiment importante ?
– Quelle est la problématique de la question posée dans l'exercice ?
– Quelles sont les méthodes mises en œuvre pour répondre à cette question ?
– Pourquoi ai-je échoué à cette question ?
– Ce que j'écris est-il cohérent ? Si ce n'est pas le cas, il faut en être conscient et ne pas laisser les choses en l'état : soit vous trouvez l'origine de l'incohérence et vous arrivez à vous corriger soit vous ne voyez pas d'où cela vient mais vous mettez clairement en avant que vous êtes conscient de cette incohérence.

N'hésitez pas à annoter vos cours, exercices et devoirs avec toutes les réponses aux questions que vous vous posez. C'est en faisant vivre vos notes que l'apprentissage se fait en profondeur.

Cet investissement actif prend du temps mais il est synonyme d'une assimilation véritable des connaissances et des pratiques donc cet investissement, à long terme, est très rentable.

Bon courage !

Je tiens à remercier chaleureusement Mme Karine Beaurpère pour ses suggestions et sa relecture minutieuse de l'ouvrage, ainsi que les Éditions Ellipses pour le soin apporté à la mise en page de cet ouvrage et pour la patience dont elles ont fait preuve.

C'est avec plaisir que je recevrai toutes vos remarques, questions, critiques et suggestions ; pour cela n'hésitez pas à m'écrire à l'adresse : `livres@tlegay.fr`.

SAVOIRS

Thème 1 - Structures algébriques usuelles

[S1.1] Magmas

- Un magma est un couple $(E,*)$, où E est un ensemble et $*$ une loi de composition interne sur E, c'est-à-dire une application de E^2 dans E.
  Si $(E,*)$ est un magma, une partie F de E est dite stable par la loi $*$ si :
  $$\forall (x,y) \in F^2,\ x * y \in F\,.$$
  La restriction de la loi $*$ à F^2 s'appelle alors la loi induite sur F.

- **Associativité**
  Soit $(E,*)$ un magma. La loi $*$ est dite associative si :
  $$\forall (x,y,z) \in E^3,\ (x * y) * z = x * (y * z)\,.$$
  Si tel est le cas, pour tout élément x de E, et tout entier n de $\mathbb{N}^*$, on pourra définir l'itéré n^{e} de x par : $x^{(n)} = \underbrace{x * x * \ldots * x}_{n \text{ fois}}$ (noté x^n si la loi est notée multiplicativement, ou nx si la loi est notée additivement).
  On a facilement, pour tous n et m de $\mathbb{N}^*$:
  $$x^{(n+m)} = x^{(n)} * x^{(m)} \quad \text{et} \quad \left(x^{(n)}\right)^{(m)} = x^{(nm)}.$$

- **Commutativité**
  Soit $(E,*)$ un magma. Deux éléments x et y de E sont dits permutables (ou qu'ils commutent) si $x * y = y * x$.
  Si x et y sont deux éléments permutables d'un magma associatif $(E,*)$, alors :
  - pour tous entiers n et m de $\mathbb{N}^*$, $x^{(n)}$ et $y^{(m)}$ sont permutables ;
  - pour tout entier $n \in \mathbb{N}^*$, $x^{(n)} * y^{(n)} = (x * y)^{(n)}$.

  La loi $*$ est dite commutative si tous les éléments commutent entre eux, soit :
  $$\forall (x,y) \in E^2,\ x * y = y * x\,.$$

- **Élément neutre**
  On appelle élément neutre d'un magma $(E,*)$ un élément $e \in E$ tel que :
  $$\forall\, x \in E,\ x * e = e * x = x\,.$$
  Si $(E,*)$ possède un élément neutre, celui-ci est unique. Un magma qui possède un élément neutre est dit unifère.

- **Symétrique**
  . Soit $(E,*)$ un magma unifère d'élément neutre e, et x un élément de E.
  On appelle symétrique à gauche (respectivement à droite) de x un élément x' (respectivement x'') de E, s'il existe, tel que :
  $$x' * x = e \quad (\text{respectivement } x * x'' = e)\,.$$
  Si x' (respectivement x'') existe, on dit que x est symétrisable à gauche (respectivement à droite).

- Soit $(E,*)$ un magma unifère d'élément neutre e, et x un élément de E.
  On appelle symétrique de x un élément x' de E, s'il existe, tel que :

$$x' * x = x * x' = e\,.$$

  Si x' existe, x est dit symétrisable.
  Cela équivaut à dire que x est symétrisable à droite *et* à gauche *et* que ses symétriques à droite et à gauche sont égaux.

  Exemple : dans l'ensemble $\mathcal{A}(E,E)$ des applications de E dans E, muni de la loi $\circ$, les éléments symétrisables à droite sont les applications surjectives et les éléments symétrisables à gauche sont les applications injectives. Les éléments symétrisables sont donc les applications bijectives de E sur E.

- Soit $(E,*)$ un magma *associatif et unifère* ($(E,*)$ s'appelle alors un *monoïde*).
  - Si un élément x de E admet un symétrique à droite et un symétrique à gauche, ceux-ci sont égaux (et x est alors symétrisable).
  - Si un élément x de E est symétrisable, son symétrique est unique.
  - Si un élément x de E est symétrisable, de symétrique x', alors x' est symétrisable, de symétrique x.
  - Si x et y sont symétrisables (de symétriques respectifs x' et y') , il en est de même de $x * y$ et : $(x * y)' = y' * x'$.

    ✓ Lorsque la loi est notée multiplicativement, on parle *d'inverse* au lieu de symétrique. Lorsqu'elle est notée additivement, on parle *d'opposé*.

- Soit $(E,*)$ un magma unifère, et x un élément de E. On note : $x^{(0)} = e$.
  Si x est symétrisable, de symétrique x', pour tout $n \in \mathbb{N}^*$ on note $x^{(-n)}$ l'élément $x'^{(n)}$ (ainsi, $x' = x^{(-1)}$).
  On peut ainsi, lorsque x est symétrisable, étendre la notation $x^{(n)}$ pour $n \in \mathbb{Z}$.
  - Soit x un élément *symétrisable* de E. Alors, pour tous n et m de $\mathbb{Z}$, on a :

$$x^{(n+m)} = x^{(n)} * x^{(m)} \quad \text{et} \quad \left(x^{(n)}\right)^{(m)} = x^{(nm)}.$$

  - Soient x et y deux éléments *permutables* et *symétrisables* de E. Alors :
    - pour tous entiers n et m de $\mathbb{Z}$, $x^{(n)}$ et $y^{(m)}$ sont permutables ;
    - pour tout entier $n \in \mathbb{Z}$, $x^{(n)} * y^{(n)} = (x * y)^{(n)}$.

- **Éléments réguliers**

  Soit $(E,*)$ un magma. Un élément a de E est dit régulier (ou simplifiable) à gauche (respectivement à droite) si :

$$\forall\,(x,y) \in E^2,\ a * x = a * y \ \Rightarrow\ x = y \text{ (respectivement } x * a = y * a \ \Rightarrow\ x = y)\,.$$

  Un élément a de E est dit régulier s'il est à la fois régulier à gauche *et* à droite.
  Soit $(E,*)$ un magma associatif et unifère. Si un élément a de E est symétrisable, (à droite, à gauche), *alors* il est régulier (à droite, à gauche).

  ✓ La réciproque de cette proposition est fausse en général.
  Par exemple, dans $(\mathbb{Z},\times)$, 2 est régulier mais non symétrisable.

[S1.2] Morphismes de magmas

- Soient $(E,*)$ et $(F,\square)$ deux magmas.

  On dit qu'une application $f : E \to F$ est un morphisme de $(E,*)$ dans $(F,\square)$ si :

  $$\forall\,(x,y) \in E^2,\ f(x*y) \ = \ f(x)\square f(y)\,.$$

  Un isomorphisme est un morphisme bijectif. Un endomorphisme est un morphisme de $(E,*)$ dans lui-même. Un automorphisme est un endomorphisme bijectif.

  Exemple : si $(E,*)$ est un magma associatif unifère et si x est un élément symétrisable de E, l'application $n \mapsto x^{(n)}$ est un morphisme de $(\mathbb{Z},+)$ dans $(E,*)$.

- Si f est un morphisme de $(E,*)$ dans $(F,\square)$ et si g un morphisme de $(F,\square)$ dans $(G,\triangle)$, la composée $g \circ f$ est un morphisme de $(E,*)$ dans $(G,\triangle)$.

- Si f est un isomorphisme de $(E,*)$ dans $(F,\square)$, alors son application réciproque f^{-1} est un isomorphisme de $(F,\square)$ dans $(E,*)$.

  Exemple : l'application $x \mapsto \mathrm{e}^x$ est un isomorphisme de $(\mathbb{R},+)$ sur $(\mathbb{R}_+^*,\times)$.

- **Transport de structure**

  Soit $f : (E,*) \to (F,\square)$ un morphisme de magmas.

  - L'image $f(E)$ de E par f est une partie stable de $(F,\square)$.
  - Si $*$ est commutative, alors $\square$ est commutative *dans le magma* $(f(E),\square)$.
  - Si $*$ est associative, alors $\square$ est associative *dans le magma* $(f(E),\square)$.
  - Si e est l'élément neutre de $(E,*)$, alors $f(e)$ est l'élément neutre de $(f(E),\square)$.
  - Si x est symétrisable dans $(E,*)$, de symétrique x', alors $f(x)$ est symétrisable dans $(f(E),\square)$, de symétrique $f(x')$, et on a alors, pour tout $n \in \mathbb{Z}$, $f\left(x^{(n)}\right) = \left(f(x)\right)^{(n)}$.

[S1.3] Groupes

- On appelle groupe un *magma* $(G,*)$ tel que :

  (i) $*$ est associative ;

  (ii) $*$ possède un élément neutre *(généralement noté e_G)* ;

  (iii) tout élément de G est symétrisable pour la loi $*$.

  Le groupe est dit abélien (ou commutatif) si, de plus, la loi $*$ est commutative.

  Exemple : soit E un ensemble non vide. Alors l'ensemble $\mathfrak{S}(E)$ des *permutations* de E (c'est-à-dire l'ensemble des bijections de E dans E) est un groupe pour la loi $\circ$; ce groupe n'est pas commutatif dès que $\mathrm{Card}(E) \geqslant 3$.

- **Produit de groupes**

  Soient $(G,*_G)$ et $(H,*_H)$ deux groupes. On peut alors munir l'ensemble produit $G \times H$ de la loi $\square$ définie par :

  $$\forall\,(x_1,y_1),(x_2,y_2) \in (G\times H)^2,\ (x_1,y_1)\square(x_2,y_2) \ = \ (x_1 *_G x_2, y_1 *_H y_2)\,.$$

  Alors $(G \times H,\square)$ est un groupe, appelé groupe produit de G et H ; son élément neutre est (e_G,e_H).

On peut bien sûr étendre cette définition à un produit d'un nombre fini quelconque de groupes.

[S1.4] Sous-groupe

- Soit $(G,*)$ un groupe. On dit qu'une *partie* H de G est un sous-groupe de G si $(H,*)$ est encore un groupe.
  Si H est un sous-groupe de G, alors :
  - l'élément neutre de H est celui de G ;
  - si x est un élément de H, son symétrique dans H est le même que dans G.

- **Caractérisation d'un sous-groupe**
  Soit $(G,*)$ un groupe. Pour qu'une *partie* H de G soit un sous-groupe de G, il faut et il suffit que les trois conditions suivantes soient vérifiées :
  (i) $H \neq \emptyset$;
  (ii) H est stable par la loi $*$;
  (iii) pour tout élément x de H, son symétrique x^{-1} est dans H.

  Ces trois conditions sont aussi équivalentes aux deux conditions suivantes :
  (i) $H \neq \emptyset$;
  (ii) $\forall (x,y) \in H^2,\ x * y^{-1} \in H$.

  Lorsque la loi est notée additivement, cette dernière condition s'écrit : $\forall (x,y) \in H^2,\ x - y \in H$.

  Exemples
  - $(\{-1,1\},\times)$ est un sous-groupe de $(\mathbb{R}^*,\times)$.
  - L'ensemble $\mathbb{U}$ des nombres complexes de module égal à 1 est un sous-groupe de $(\mathbb{C}^*,\times)$ (il s'agit du *cercle unité*).
  - Si n est un entier non nul, l'ensemble $\mathbb{U}_n$ des racines n-ièmes de l'unité est un sous-groupe de $(\mathbb{U},\times)$.

- **Théorème de Lagrange**
  Si G est un groupe fini, le cardinal de tout sous-groupe de G est un diviseur de $\operatorname{Card} G$.

- **Intersection de sous-groupes**
  Soit $(G,*)$ un groupe. L'intersection d'une famille $(H_i)_{i\in I}$ de sous-groupes de G est encore un sous-groupe de G.

  ✓ La *réunion* de sous-groupes de G n'est pas en général un sous-groupe de G. Plus précisément, si H et H' sont deux sous-groupes de G, $H \cup H'$ est encore un sous-groupe de G si et seulement si $H \subset H'$ ou $H' \subset H$.

- **Sous-groupe engendré**
  - Soit $(G,*)$ un groupe, et X une partie de G. L'intersection de tous les sous-groupes de G contenant X est un sous-groupe de G ; c'est le plus petit sous-groupe de G contenant X (au sens de l'inclusion) ; on l'appelle sous-groupe engendré par X, et on le note $\operatorname{gr}(X)$.

- Si $X = \emptyset$, $\mathrm{gr}(\emptyset) = \{e_{_G}\}$. Sinon, $\mathrm{gr}(X)$ est exactement l'ensemble des éléments de la forme :
$$x_1 * x_2 * \cdots * x_n$$
où $n \in \mathbb{N}^*$ et où pour tout $i \in [\![1\,;n]\!]$, $x_i \in X$ ou $x_i^{-1} \in X$.

[S1.5] Morphismes de groupes

- Un morphisme de groupes est (tout simplement) un morphisme entre deux groupes $(G,*)$ et $(H,\square)$.
  On définit de la même façon qu'auparavant les notions d'iso-, d'endo- et d'automorphisme de groupes.
- **Propriétés**
  - Soit f un morphisme d'un groupe G vers un groupe H. Alors :
    - $f(e_{_G}) = e_{_H}$;
    - $\forall x \in G,\ \forall n \in \mathbb{Z},\ f\left(x^{(n)}\right) = (f(x))^{(n)}$.
  - La composée de deux morphismes de groupes est un morphisme de groupes.
  - Si f est un isomorphisme de groupes, il en est de même de f^{-1}.
  - Si G est un groupe, l'ensemble $\mathrm{Aut}(G)$ des automorphismes de G est un groupe pour la loi $\circ$; c'est un sous-groupe du groupe des permutations $\big(\mathfrak{S}(G),\circ\big)$.
- **Images directe et réciproque d'un sous-groupe**
  - Soit f un morphisme d'un groupe $(G,*)$ vers un groupe $(H,\square)$.
    Si G' est un sous-groupe de $(G,*)$, son image $f(G')$ est un sous-groupe de $(H,\square)$.
  - Si H' est un sous-groupe de $(H,\square)$, son image réciproque par f, $f^{-1}(H')$, est un sous-groupe de $(G,*)$.
- Si f est un morphisme d'un groupe $(G,*)$ vers un groupe $(H,\square)$, on appelle :
  - image de f, notée $\mathrm{Im}\, f$, l'image de G par f, soit : $\mathrm{Im}\, f = \{f(x)\ ,\ x \in G\}$;
  - noyau de f, noté $\mathrm{Ker}\, f$, l'image réciproque de $\{e_{_H}\}$ par f, soit :
$$\mathrm{Ker}\, f = \{x \in G \mid f(x) = e_{_H}\}\,.$$
  On alors les résultats suivants.
  - $\mathrm{Im}\, f$ est un sous-groupe de H, et : f surjective $\iff \mathrm{Im}\, f = H$.
  - $\mathrm{Ker}\, f$ est un sous-groupe de G, et : f injective $\iff \mathrm{Ker}\, f = \{e_{_G}\}$.
- **Exemples**
  - L'application « déterminant » est un morphisme du magma $(\mathcal{M}_n(\mathbb{K}),\times)$ dans le magma $(\mathbb{R},\times)$.
  - C'est aussi un morphisme du groupe $(\mathrm{GL}_n(\mathbb{K}),\times)$ dans le groupe $(\mathbb{R}^*,\times)$.
  - C'est aussi un morphisme du groupe orthogonal $(\mathcal{O}_n(\mathbb{R}),\times)$ dans $(\{-1,1\},\times)$, sous-groupe de $(\mathbb{R}^*,\times)$.
    Dans ce cas, son noyau est l'ensemble des matrices orthogonales de déterminant $+1$; c'est le groupe spécial orthogonal $\mathcal{O}_n^+(\mathbb{R})$.

[S1.6] Groupes monogènes et cycliques

- **Sous-groupes de** $(\mathbb{Z},+)$

  Les sous-groupes de $(\mathbb{Z},+)$ sont exactement les ensembles $n\mathbb{Z}$ où $n \in \mathbb{N}$ (ensemble des multiples de n).

  $(n\mathbb{Z},+)$ est le sous-groupe de $(\mathbb{Z},+)$ engendré par n ou par $-n$.

- **Le groupe** $(\mathbb{Z}/n\mathbb{Z},+)$

  . **Congruences**

  Soit $n \in \mathbb{N}$. On dit que deux entiers x, $y \in \mathbb{Z}$ sont congrus modulo n, et l'on note $x \equiv y \pmod{n}$, s'il existe $k \in \mathbb{Z}$ tel que $y = x + kn$.

  Cela équivaut à dire que $x - y$ appartient à $n\mathbb{Z}$, ou que (lorsque $n \geqslant 1$), x et y ont le même reste dans la division euclidienne par n.

  Il s'agit d'une relation d'équivalence sur $\mathbb{Z}$, compatible avec l'addition et la multiplication, c'est-à-dire que pour tous entiers x, y, x', y' :

  $$x \equiv y \pmod{n} \text{ et } x' \equiv y' \pmod{n} \\ \Longrightarrow x + x' \equiv y + y' \pmod{n} \text{ et } xx' \equiv yy' \pmod{n}.$$

  En particulier, pour tout $(x,y) \in \mathbb{Z}^2$ et tout entier naturel $k \in \mathbb{N}$:

  $$x \equiv y \pmod{n} \Longrightarrow x^k \equiv y^k \pmod{n}.$$

  . La classe d'équivalence de x modulo n est l'ensemble des $y \in \mathbb{Z}$ congrus à x :

  $$\overline{x} = \{y \in \mathbb{Z} \mid x \equiv y \pmod{n}\} = \{x\} + n\mathbb{Z}.$$

  L'ensemble de toutes les classes d'équivalence modulo n se note $\mathbb{Z}/n\mathbb{Z}$.
  Si $n = 0$, $\mathbb{Z}/0\mathbb{Z} = \mathbb{Z}$; si $n = 1$, $\mathbb{Z}/\mathbb{Z} = \{\overline{0}\}$; sinon, pour $n \geqslant 2$ (ce que l'on supposera pour la suite),

  $$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

  . **Addition dans** $\mathbb{Z}/n\mathbb{Z}$

  Elle est définie par :

  $$\forall (\overline{x},\overline{y}) \in (\mathbb{Z}/n\mathbb{Z})^2,\ \overline{x} + \overline{y} = \overline{x+y}.$$

  Muni de cette loi, l'ensemble $(\mathbb{Z}/n\mathbb{Z},+)$ est un groupe abélien ; l'application $\pi : \begin{cases} \mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \\ x & \longmapsto & \overline{x} \end{cases}$ est alors un morphisme de groupes surjectif, dont le noyau est $n\mathbb{Z}$. π s'appelle la surjection canonique de $\mathbb{Z}$ sur $\mathbb{Z}/n\mathbb{Z}$.

- **Groupe monogène, cyclique**

  . Un groupe $(G,*)$ est dit monogène s'il est engendré par un unique élément $a \in G$. Dans ce cas, en notant la loi multiplicativement :

  $$G = \{a^k,\ k \in \mathbb{Z}\}.$$

  Un groupe est dit cyclique s'il est monogène et de cardinal fini.

- Tout groupe monogène de cardinal infini est isomorphe à $(\mathbb{Z},+)$.
- Si G est un groupe cyclique engendré par a et de cardinal $n \in \mathbb{N}^*$, G est isomorphe au groupe $(\mathbb{Z}/n\mathbb{Z},+)$. De plus, on a alors : $a^n = e_G$.

- **Générateurs d'un groupe cyclique**

  Soit $n \geqslant 2$. Les générateurs du groupe $(\mathbb{Z}/n\mathbb{Z},+)$ sont les éléments $\overline{k}$ où k est premier avec n.

  Le nombre de ces générateurs, c'est-à-dire le nombre d'entiers de $[\![0\,;n-1]\!]$ premiers avec n se note $\varphi(n)$; φ s'appelle l'indicateur d'Euler.

  Par isomorphisme, il en résulte que si $(G,*)$ est un groupe cyclique engendré par a et de cardinal n, les générateurs de G sont exactement les éléments a^k où $k \in \mathbb{Z}$ est un entier premier avec n.

  Exemple : l'ensemble $\mathbb{U}_n$ des racines n^{es} de l'unité est un sous-groupe cyclique de $(\mathbb{C}^*,\times)$.

  Ses générateurs sont les nombres complexes de la forme $\mathrm{e}^{\frac{2ik\pi}{n}}$ avec $k \in [\![0\,;n-1]\!]$ premier avec n. Ces nombres sont appelés les racines primitives n^{es} de l'unité.

[S1.7] Ordre d'un élément dans un groupe

- Un élément a d'un groupe $(G,*)$ est dit d'ordre fini si le groupe engendré par a,

$$\mathrm{gr}(a) = \left\{a^k,\, k \in \mathbb{Z}\right\},$$

  est de cardinal fini d. d s'appelle alors l'ordre de a.

  Si a est d'ordre fini, l'ordre de a est le plus petit entier $n \in \mathbb{N}^*$ tel que $a^n = e_G$; plus précisément, si d est l'ordre de a :

$$a^n = e_G \iff n \text{ multiple de } d\,.$$

- D'après le théorème de Lagrange, si $(G,*)$ est un groupe fini de cardinal n, tout élément $a \in G$ est d'ordre fini et son ordre d divise le cardinal de G $(d \,|\, n)$.

  En particulier, pour tout $a \in G$, on a $a^n = e_G$.

- En regroupant alors les éléments du groupe $(\mathbb{Z}/n\mathbb{Z},+)$ selon leur ordre, on obtient la propriété suivante de l'indicateur d'Euler φ :

$$n = \sum_{d \text{ divise } n} \varphi(d)\,.$$

[S1.8] Le groupe symétrique $\mathfrak{S}_n$

- Soit $n \in \mathbb{N}^*$. On appelle groupe symétrique d'ordre n, noté $\mathfrak{S}_n$, l'ensemble des permutations de l'ensemble $[\![1\,;n]\!]$.

  $\mathfrak{S}_n$ est un groupe pour la loi $\circ$ de composition des applications. C'est un groupe fini de cardinal $n!$, non commutatif dès que $n \geqslant 3$.

- **Cycles, transpositions**
  - Soit $n \geqslant 2$ et $p \in [\![2\,;n]\!]$. On appelle cycle de longueur p toute permutation $\sigma \in \mathfrak{S}_n$ telle qu'il existe p éléments distincts $a_1,a_2,\ldots,a_p$ de $[\![1\,;n]\!]$ tels que :
$$\sigma(a_1) = a_2\,,\ \sigma(a_2) = a_3\,,\ \ldots\,,\ \sigma(a_{p-1}) = a_p\,,\ \sigma(a_p) = a_1$$
$$\text{et}\ \ \forall\, x \notin \{a_1,\ldots,a_p\}\,,\ \sigma(x) = x\,.$$
L'ensemble $\{a_1,a_2,\ldots,a_p\}$ s'appelle le support du cycle.
  - Un cycle de longueur p est un élément d'ordre p du groupe $(\mathfrak{S}_n,\circ)$.
  - Un cycle de longueur 2 s'appelle une transposition.
- **Générateurs du groupe** $(\mathfrak{S}_n,\circ)$
  - Toute permutation de $\mathfrak{S}_n$ peut s'écrire comme composée commutative de cycles à supports disjoints. Cette décomposition est unique à l'ordre près des facteurs.
  - Toute permutation de $\mathfrak{S}_n$ peut s'écrire comme composée (non commutative et non unique) de transpositions.
- **Signature d'une permutation**
  - Soit $\sigma \in \mathfrak{S}_n$ $(n \geqslant 2)$. On appelle inversion de σ tout couple $(i,j) \in [\![1\,;n]\!]^2$ tel que : $i < j$ et $\sigma(i) > \sigma(j)$.
On appelle alors signature de σ le nombre $\varepsilon(\sigma) = (-1)^{\nu(\sigma)}$ où $\nu(\sigma)$ est le nombre d'inversions dans σ. Une permutation est dite paire si $\varepsilon(\sigma) = 1$ et impaire sinon.
  - Si σ est un cycle de longueur p, $\varepsilon(\sigma) = (-1)^{p-1}$. En particulier, la signature d'une transposition est égale à -1.
  - La fonction ε est un morphisme du groupe $(\mathfrak{S}_n,\circ)$ sur le groupe $(\{-1,1\},\times)$:
$$\forall\, \sigma,\, \sigma' \in \mathfrak{S}_n,\ \varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma)\varepsilon(\sigma')\,.$$
Son noyau, c'est-à-dire l'ensemble des permutations paires, est un sous-groupe de $(\mathfrak{S}_n,\circ)$, noté $\mathcal{A}_n$ et appelé groupe alterné d'ordre n.

[S1.9] Anneaux

- Un anneau $(A,+,\times)$ est un triplet formé d'un ensemble A, de deux lois de compositions internes $+$ et $\times$, tel que :
  (i) $(A,+)$ est un groupe abélien (dont l'élément neutre sera noté 0_A) ;
  (ii) $\times$ est associative et possède un élément neutre (noté 1_A et aussi appelé élément unité) ;
  (iii) $\times$ est distributive à gauche *et* à droite par rapport à $+$, c'est-à-dire :
$$\forall\,(x,y,z) \in A^3,\ x(y+z) = xy + xz \text{ et } (y+z)x = yx + zx\,.$$
L'anneau est dit commutatif si la loi $\times$ est commutative.

> ✓ Une loi peut être distributive à droite *mais pas* à gauche ; c'est le cas par exemple des lois $+$ et $\circ$ dans $\mathcal{A}(\mathbb{R},\mathbb{R})$

- **Règles de calcul dans un anneau**

  Soit $(A, +, \times)$ un anneau.

  - $\forall x \in A,\ 0_A \times x = x \times 0_A = 0_A$.
  - $\forall (x,y) \in A^2,\ x(-y) = (-x)y = -(x \times y)$ et $(-x)(-y) = xy$.
  - $\forall (x,y,z) \in A^3,\ x(y-z) = xy - xz$ et $(y-z)x = yx - zx$.
  - Si x et y sont deux éléments *permutables* de A, alors, pour tout $n \in \mathbb{N}$:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \quad \text{(formule du binôme)}.$$

  - Si x et y sont deux éléments *permutables* de A, alors, pour tout $n \in \mathbb{N}^*$:

$$x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1}).$$

  En particulier, pour tout $x \in A$:

$$1_A - x^n = (1_A - x)(1_A + x + \cdots + x^{n-1}).$$

- **Groupe des éléments inversibles**

  Soit $(A, +, \times)$ un anneau.

  L'ensemble $\mathcal{U}(A)$ des éléments inversibles (pour la loi $\times$) de l'anneau A est un groupe pour la loi $\times$ (d'élément neutre 1_A).

- **Sous-anneaux**

  . Soit $(A, +, \times)$ un anneau, et B une *partie* de A.

  On dit que B est un sous-anneau de A si $(B, +, \times)$ est encore un anneau, *de même élément unité* que A.

  . **Caractérisation d'un sous-anneau**

  Soit $(A,+,\times)$ un anneau, et B une *partie* de A. Pour que B soit un sous-anneau de A, il faut et il suffit que les trois conditions suivantes soient vérifiées :

  (i) $\forall (x,y) \in B^2,\ x - y \in B$;

  (ii) $\forall (x,y) \in B^2,\ xy \in B$;

  (iii) $1_A \in B$.

- **Anneau produit**

  Si $(A, +, \cdot)$ et $(B, +, \cdot)$ sont deux anneaux, l'ensemble $A \times B$ est muni d'une structure d'anneau par les lois définies, pour tous (a,b), (a',b') de $A \times B$, par :

$$(a,b) + (a',b') = (a+a', b+b') \quad \text{et} \quad (a,b) \times (a',b') = (a \cdot a', b \cdot b').$$

- **Morphismes d'anneaux**

  . Soient A et A' deux anneaux (dont les lois sont notées de façon identique).

  Une application f de A dans A' est appelée un morphisme d'anneaux si :

  (i) $f(1_A) = 1_{A'}$;

  (ii) $\forall (x,y) \in A^2,\ f(x+y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$.

. Soit $f : A \to A'$ un morphisme d'anneaux.

– Si B est un sous-anneau de A, $f(B)$ est un sous-anneau de A'.

– Si B' est un sous-anneau de A', son image réciproque $f^{-1}(B')$ est un sous-anneau de A.

- **Anneau intègre**

. Un anneau $(A, +, \times)$ est dit intègre s'il n'est pas réduit à $\{0\}$ et si :

$$\forall\,(a,b) \in A^2,\ ab = 0 \implies a = 0 \text{ ou } b = 0\,.$$

Ainsi, un anneau A est intègre si et seulement si le produit de deux éléments de A différents de 0_A est différent de 0_A.
Dans un anneau intègre, tous les éléments sont réguliers pour la loi $\times$.

. **Exemples**

– Si $\mathbb{K}$ est un corps, l'anneau $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ est intègre.

– Pour $n \geqslant 2$, l'anneau des matrices carrées $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre.

[S1.10] Idéaux d'un anneau commutatif

- Soit $(A,+,\times)$ un anneau *commutatif*. Une partie I de A est un idéal de A si c'est un sous-groupe additif et s'il est absorbant pour la loi $\times$, c'est-à-dire :

(i) $(I,+)$ est un sous-groupe de $(A,+)$;

(ii) $\forall\, a \in A,\ \forall\, x \in I,\ ax \in I$.

- **Caractérisation d'un idéal**
Une partie I d'un anneau commutatif $(A,+,\times)$ est un idéal de A si et seulement si :

(i) $I \neq \emptyset$;

(ii) $\forall\,(x,y) \in I^2,\ x + y \in I$;

(iii) $\forall\, a \in A,\ \forall\, x \in I,\ ax \in I$.

- L'intersection d'une famille quelconque d'idéaux d'un anneau commutatif A est encore un idéal de A.

- Si $I_1, \ldots, I_p$ sont des idéaux d'un d'un anneau commutatif A, l'ensemble :

$$I_1 + \cdots + I_p = \{x_1 + \cdots + x_p \mid \forall\, k \in [\![1\,;p]\!],\ x_k \in I_k\}$$

est encore un idéal de A.

- **Idéal engendré**

. Soit X une partie quelconque d'un anneau commutatif A.
L'intersection de tous les idéaux de A qui contiennent X est encore un idéal de A ; c'est le plus petit idéal de A contenant X (au sens de l'inclusion) ; on l'appelle idéal engendré par X, et on le note $\mathrm{id}(X)$.

- Soit X une partie non vide d'un anneau commutatif A.
  $\mathrm{id}(X)$ est alors l'ensemble de tous les éléments de la forme :

$$a_1x_1 + \dots a_nx_n \quad \text{avec } n \in \mathbb{N}^*,\ a_i \in A\,,\ x_i \in X.$$

- En particulier, l'idéal engendré par un élément $\{x\}$ est l'ensemble :

$$\mathrm{id}(\{x\}) = \{ax,\ a \in A\}\,.$$

  On le note aussi : Ax, ou xA (ces deux écritures désignent le même ensemble puisque l'anneau A est supposé commutatif ; sinon, il faudrait distinguer les notions *d'idéal à gauche* et *d'idéal à droite*).

- **Exemples dans l'anneau $\mathbb{Z}$**
  - Si $n \in \mathbb{N}$, $\mathrm{id}(\{n\}) = n\mathbb{Z}$.
  - Soient $a,\ b \in \mathbb{Z}$;

$$\mathrm{id}(\{a,b\}) = a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}, \text{ avec } d = \mathrm{pgcd}(a,b) \text{ (noté } a \wedge b)\,;$$
$$a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}, \text{ avec } m = \mathrm{ppcm}(a,b) \text{ (noté } a \vee b)\,.$$

- Un idéal I d'un anneau commutatif A est dit principal s'il est engendré par un seul élément (c'est-à-dire s'il existe $x \in A$ tel que $I = Ax$).
  Un anneau commutatif A est dit principal s'il est intègre et si tous ses idéaux sont principaux.

  Exemples
  - L'anneau $\mathbb{Z}$ est principal : les idéaux de l'anneau $\mathbb{Z}$ sont les ensembles $n\mathbb{Z}$, $n \in \mathbb{N}$.
  - Si $\mathbb{K}$ est un corps, l'anneau $\mathbb{K}[X]$ est principal (voir ci-après).

- **Morphismes d'anneaux**

Soient A et A' deux anneaux commutatifs (dont les lois sont notées de la même façon), et $f : A \to A'$ un morphisme d'anneaux.

(i) Si I est un idéal de A, son image $f(I)$ est un idéal du sous-anneau $f(A)$.

(ii) Si I' est un idéal de A', son image réciproque $f^{-1}(I')$ est un idéal de A.

En particulier, le noyau de f :

$$\mathrm{Ker}\, f = \{x \in A \mid f(x) = 0_{A'}\}\,,$$

est un idéal de A.

[S1.11] L'anneau $\mathbb{Z}/n\mathbb{Z}$ $(n \geqslant 2)$

- La congruence modulo n étant compatible avec la multiplication dans $\mathbb{Z}$, on peut définir dans l'ensemble quotient $\mathbb{Z}/n\mathbb{Z}$ une multiplication interne par :

$$\forall\, (\overline{x},\overline{y}) \in (\mathbb{Z}/n\mathbb{Z})^2,\ \overline{x} \cdot \overline{y} = \overline{xy}\,.$$

L'ensemble $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif.

- **Éléments inversibles**

  Les éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$ sont les $\overline{x}$ avec x premier avec n.

  Leur nombre est égal à $\varphi(n)$, où φ est l'indicateur d'Euler. Le groupe des éléments inversibles $\big(\mathcal{U}(\mathbb{Z}/n\mathbb{Z}),\times\big)$ étant de cardinal $\varphi(n)$, on a :

$$\forall\, a \in \mathbb{Z},\ a \wedge n = 1 \implies a^{\varphi(n)} \equiv 1 \ (\text{mod } n)$$

  (c'est le *théorème d'Euler*).

- **Théorème des restes chinois**

  - Si $x \in \mathbb{Z}$, nous noterons $\overline{x}_{(n)}$ sa classe d'équivalence modulo n.

    Si n et m sont deux entiers premiers entre eux, l'application :

$$\phi : \begin{cases} \mathbb{Z}/nm\mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \\ \overline{x}_{(nm)} & \longmapsto & \big(\overline{x}_{(n)},\overline{x}_{(m)}\big) \end{cases}$$

    est un isomorphisme d'anneaux.

  - **Conséquence : système de congruences**

    Soient n et m deux entiers premiers entre eux, et a, $b \in \mathbb{Z}$. L'ensemble des solutions du système :

$$\begin{cases} x \equiv a \ (\text{mod } n) \\ x \equiv b \ (\text{mod } m) \end{cases}$$

    est de la forme $\mathscr{S} = \{x_0 + kmn,\ k \in \mathbb{Z}\}$ où x_0 est une solution particulière appartenant à $[\![0\,;mn-1]\!]$.

  - **Calcul de l'indicateur d'Euler**

    La restriction de l'isomorphisme ϕ ci-dessus est un isomorphisme entre le groupe $\big(\mathcal{U}(\mathbb{Z}/nm\mathbb{Z}),\times\big)$ et le groupe produit $\big(\mathcal{U}(\mathbb{Z}/n\mathbb{Z}),\times\big) \times \big(\mathcal{U}(\mathbb{Z}/m\mathbb{Z}),\times\big)$ des éléments inversibles.

    Il en résulte que, si φ désigne l'indicateur d'Euler :

$$n \wedge m = 1 \implies \varphi(nm) = \varphi(n)\varphi(m)\,,$$

    puis que, si $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ est la décomposition de n en facteurs premiers :

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \times \cdots \times \left(1 - \frac{1}{p_r}\right).$$

[S1.12] L'anneau $\mathbb{K}[X]$

$\mathbb{K}$ désigne ici un sous-corps de $\mathbb{C}$. Les résultats de 1re année concernant les polynômes à coefficients réels ou complexes s'étendent au cas des polynômes à coefficients dans $\mathbb{K}$. Nous rappelons brièvement les principaux résultats.

- **Construction**

  - Soit $\mathbb{K}$ un sous-corps de $\mathbb{C}$. On appelle polynôme à coefficients dans $\mathbb{K}$ toute suite $P = (a_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{K}$ qui est nulle à partir d'un certain rang.

    Leur ensemble est noté $\mathbb{K}[X]$.

  - Pour les lois usuelles sur les suites (addition et multiplication externe), $\mathbb{K}[X]$ est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$.

    On définit de plus une multiplication interne dans $\mathbb{K}[X]$ de la façon suivante :

si $P = (a_n)_{n\in\mathbb{N}}$ et $Q = (b_n)_{n\in\mathbb{N}}$, alors $PQ = (c_n)_{n\in\mathbb{N}}$ où :

$$\forall n \in \mathbb{N},\ c_n = \sum_{k=0}^{n} a_k b_{n-k}\,.$$

- Si l'on note 1 le polynôme $1 = (1,0,0,\ldots)$ et X le polynôme $X = (0,1,0,0,\ldots)$, alors tout polynôme $P \in \mathbb{K}[X]$ s'écrit de façon unique sous la forme :

$$P = \sum_{k=0}^{+\infty} a_k X^k\,,$$

où les a_k sont nuls à partir d'un certain rang.
Lorsque P n'est pas le polynôme nul, on peut définir son degré :

$$\deg P = \max\{k \in \mathbb{N} \mid a_k \neq 0\}\,.$$

Par convention, le degré du polynôme nul est $-\infty$.
Un polynôme est dit unitaire s'il est non nul et son coefficient dominant est égal à 1 ; un polynôme est dit normalisé s'il est nul ou unitaire.

- $(\mathbb{K}[X],+,\times)$ est alors un anneau commutatif et intègre. Ses éléments inversibles sont les polynômes constants non nuls.

- **Division euclidienne dans $\mathbb{K}[X]$**
Soient $A,\ B \in \mathbb{K}[X]$ avec $B \neq 0$. Il existe un et un seul couple $(Q,R) \in \mathbb{K}[X]^2$ tel que :

$$A = BQ + R \text{ avec } \deg R < \deg B\,.$$

- **Idéaux de $\mathbb{K}[X]$**
Tout idéal de l'anneau $\mathbb{K}[X]$ est principal. Plus précisément, si I est un idéal de $\mathbb{K}[X]$, il existe un et un seul polynôme normalisé A tel que :

$$I = A.\mathbb{K}[X] = \{AQ \mid Q \in \mathbb{K}[X]\}\,.$$

- **Racines d'un polynôme**
Si $P = \sum_{k=0}^{d} a_k X^k$ est un polynôme à coefficients dans $\mathbb{K}$, on peut lui associer la fonction polynôme (encore notée P par abus de notation) :

$$P\colon x \in \mathbb{K} \longmapsto \sum_{k=0}^{d} a_k x^k.$$

Un élément $a \in \mathbb{K}$ s'appelle une racine de P si $P(a) = 0$, ce qui équivaut à dire que le polynôme $X - a$ divise P dans $\mathbb{K}[X]$.
Plus généralement, a est appelé racine de P d'ordre $k \in \mathbb{N}$ si $(X-a)^k$ divise P et $(X-a)^{k+1}$ ne divise pas P.

- **Formule de Taylor**
Soit $P \in \mathbb{K}[X]$, de degré n, et soit $a \in \mathbb{K}$. Alors :

$$P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k.$$

Conséquence : $a \in \mathbb{K}$ est racine d'ordre $k \in \mathbb{N}^*$ de P si et seulement si :

$$P(a) = \cdots = P^{(k-1)}(a) = 0 \quad \text{et} \quad P^{(k)}(a) \neq 0\,.$$

- **Relations coefficients-racines**

  Soit $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$ un polynôme scindé de degré $n \in \mathbb{N}^*$. Notons $x_1,\ldots,x_n$ ses racines, distinctes ou non.

  On définit alors les fonctions symétriques élémentaires $\sigma_1,\ldots,\sigma_n$ des racines de P par :

  $$\forall\, k \in [\![1\,;n]\!],\ \sigma_k = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} x_{i_1}x_{i_2}\cdots x_{i_k}\,.$$

  En particulier, σ_1 désigne la somme des n racines et σ_n leur produit.

  On a alors les relations :

  $$\forall\, k \in [\![1\,;n]\!],\ \sigma_k = (-1)^k \frac{a_{n-k}}{a_n}\,.$$

- **Polynômes d'interpolation de Lagrange**

  Soit $n \in \mathbb{N}$ et soient $a_0,a_1,\ldots,a_n$ $n+1$ scalaires *deux à deux distincts.*

  Pour tout $(b_0,b_1,\ldots,b_n) \in \mathbb{K}^{n+1}$, il existe un et un seul polynôme P de degré inférieur ou égal à n tel que $P(a_i) = b_i$ pour tout $i \in [\![0\,;n]\!]$.

  P s'appelle le polynôme interpolateur de Lagrange relatif aux points (a_i,b_i)

  Si l'on note, pour tout $j \in [\![0\,;n]\!]$:

  $$L_j = \prod_{\substack{i=0\\ i\neq j}}^{n} \left(\frac{X-a_i}{a_j-a_i}\right),$$

  on a les propriétés suivantes :

  – $\forall\,(i,j) \in [\![0\,;n]\!]^2,\ L_j(a_i) = \delta_{ij}$;

  – les polynômes $(L_j)_{0\leqslant j\leqslant n}$ forment une base de $\mathbb{K}_n[X]$;

  – l'unique polynôme P de $\mathbb{K}_n[X]$ qui vérifie $P(a_i) = b_i$ pour tout $i \in [\![0\,;n]\!]$ est le polynôme $P = \sum_{j=0}^{n} b_j L_j$.

[S1.13] Arithmétique dans $\mathbb{K}[X]$

Le fait que $\mathbb{K}[X]$ soit un anneau principal permet d'étendre à $\mathbb{K}[X]$ les notions d'arithmétique dans $\mathbb{Z}$ vues en 1re année.

- **Divisibilité dans $\mathbb{K}[X]$**

  . On dit qu'un polynôme A divise un polynôme B (ou que B est un multiple de A) si et seulement si il existe $Q \in \mathbb{K}[X]$ tel que $B = AQ$.

  On note alors : $A \mid B$.

  Cela équivaut à dire que l'idéal $\mathbb{K}[X].B$ engendré par B est inclus dans l'idéal $\mathbb{K}[X].A$ engendré par A.

- On dit que A et B sont associés si A divise B et si B divise A, ce qui équivaut à dire qu'il existe un scalaire λ non nul tel que $B = \lambda A$.
  En particulier, tout polynôme est associé à un polynôme normalisé et un seul.

- **PGCD et PPCM**

  - **PGCD de deux polynômes**
    Soient A, $B \in \mathbb{K}[X]$. L'idéal engendré par A et B est l'ensemble :
    $$A.\mathbb{K}[X] + B.\mathbb{K}[X] = \left\{ AU + BV \;\middle|\; (U,V) \in \mathbb{K}[X]^2 \right\}.$$
    $\mathbb{K}[X]$ étant principal, il existe un polynôme normalisé D et un seul tel que $A.\mathbb{K}[X] + B.\mathbb{K}[X] = D.\mathbb{K}[X]$.
    Par définition, D s'appelle le pgcd de A et B ; il est noté $A \wedge B$.
    Ainsi, si $D = A \wedge B$, il existe U et $V \in \mathbb{K}[X]$ tels que $AU + BV = D$ (*identité de Bezout*), et réciproquement, tout polynôme de la forme $AU + BV$ est un multiple de D.

  - **Caractérisation du pgcd**
    D est le pgcd de A et de B si et seulement si :
    (i) D est normalisé ;
    (ii) $D \big| A$ et $D \big| B$;
    (iii) $\forall P \in \mathbb{K}[X]$, $P \big| A$ et $P \big| B \implies P \big| D$.

  - **PPCM de deux polynômes**
    Soient A, $B \in \mathbb{K}[X]$. Le ppcm de A et B est l'unique polynôme normalisé M qui engendre l'idéal $A.\mathbb{K}[X] \cap B.\mathbb{K}[X]$:
    $$A.\mathbb{K}[X] \cap B.\mathbb{K}[X] = M.\mathbb{K}[X].$$
    On note $M = A \vee B$.

  - **Caractérisation du ppcm**
    M est le ppcm de A et de B si et seulement si :
    (i) M est normalisé ;
    (ii) $A \big| M$ et $B \big| M$;
    (iii) $\forall P \in \mathbb{K}[X]$, $A \big| P$ et $B \big| P \implies M \big| P$.

  - **Généralisation**
    Si $A_1, \ldots, A_n$ sont des polynômes, on peut définir leur pgcd D et leur ppcm M comme étant les polynômes normalisés tels que :
    $$\sum_{i=1}^{n} A_i.\mathbb{K}[X] = D.\mathbb{K}[X] \quad \text{et} \quad \bigcap_{i=1}^{n} A_i.\mathbb{K}[X] = M.\mathbb{K}[X].$$
    Les lois « pgcd » et « ppcm » sont associatives ; autrement dit, on peut calculer les pgcd et ppcm de n polynômes en calculant de proche en proche :
    $$\bigwedge_{i=1}^{n} A_i = \left(\bigwedge_{i=1}^{n-1} A_i \right) \wedge A_n \quad \text{et} \quad \bigvee_{i=1}^{n} A_i = \left(\bigvee_{i=1}^{n-1} A_i \right) \vee A_n .$$

- **Polynômes premiers entre eux**

  . Deux polynômes A et B de $\mathbb{K}[X]$ sont dits premiers entre eux si $A \wedge B = 1$.
  Exemple : les polynômes $X - a$ et $X - b$ sont premiers entre eux si et seulement si a est différent de b.

  . **Théorème de Bezout**
  Deux polynômes A et B sont premiers entre eux si et seulement si il existe deux polynômes U et V tels que : $AU + BV = 1$.

  . **Théorème de Gauss**
  Soient A, B, $C \in \mathbb{K}[X]$.
  Si A divise le produit BC et est premier avec B, alors A divise C.

  . **Propriétés**

  – $\big(A \wedge B = 1 \text{ et } A \wedge C = 1\big) \iff A \wedge BC = 1$.

  – On en déduit, par récurrence :

    – si A est premier avec B_1,B_2,…,B_n, il est premier avec leur produit ;

    – si $A \wedge B = 1$, alors $A^n \wedge B^m = 1$ pour tout $(n,m) \in \mathbb{N}^2$.

  – Si A est divisible par B et divisible par C, et si B et C sont premiers entre eux, alors A est divisible par le produit BC (ce résultat se généralise au cas où A est divisible par n polynômes premiers entre eux deux à deux).
  Exemple : si $a_1,\ldots,a_n$ sont n racines deux à deux distinctes d'un polynôme P, alors le produit $(X - a_1)\cdots(X - a_n)$ divise P.

  . **Généralisation**
  Des polynômes $A_1,\ldots,A_n$ sont dits premiers entre eux dans leur ensemble si leur pgcd est égal à 1.
  Cela équivaut à dire qu'il existe des polynômes $U_1,\ldots,U_n$ tels que $\sum_{i=1}^{n} A_i U_i = 1$ (théorème de Bezout).
  Si n polynômes $A_1,\ldots,A_n$ sont dits premiers entre eux deux à deux, ils sont premiers entre eux dans leur ensemble.

  ✓ La réciproque de cette propriété est fausse : considérer par exemple les polynômes $(X-1)(X-2)$, $(X-2)(X-3)$ et $(X-1)(X-3)$.

- **Polynômes irréductibles**

  . Un polynôme $P \in \mathbb{K}[X]$ est dit irréductible si et seulement si :

  (i) P n'est pas constant ;

  (ii) les seuls diviseurs de P dans $\mathbb{K}[X]$ sont les polynômes constants et les polynômes associés à P.

  Exemples :

  – si $a \in \mathbb{K}$, le polynôme $X - a$ est irréductible dans $\mathbb{K}[X]$;

  – le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$ mais pas dans $\mathbb{R}[X]$;

  – le polynôme $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$ mais pas dans $\mathbb{C}[X]$.

. **Propriétés**

– Soit P un polynôme irréductible de $\mathbb{K}[X]$ et $A \in \mathbb{K}[X]$.
Alors A et P sont premiers entre eux si et seulement si A n'est pas multiple de P.

– Si P et Q sont deux polynômes irréductibles, P et Q sont sont premiers entre eux si et seulement si ils ne sont pas associés.

– Si P est irréductible et divise le produit AB, alors il divise A ou il divise B (ce résultat se généralise au cas où P divise un produit $B_1B_2\cdots B_n$).

. **Décomposition en produit de polynômes irréductibles**
Soit P un polynôme non constant de $\mathbb{K}[X]$.
Alors P peut s'écrire de manière unique (à l'ordre près des facteurs) sous la forme :

$$P = \lambda P_1^{\alpha_1} \cdots P_r^{\alpha_r},$$

où λ est un scalaire non nul, $P_1, \ldots, P_r$ sont des polynômes irréductibles deux à deux distincts et les α_i sont des entiers strictement positifs.
α_i s'appelle la valuation de P_i dans P.

. **Décomposition en produit de polynômes irréductibles dans $\mathbb{C}[X]$**
Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.
Tout polynôme P non constant de $\mathbb{C}[X]$ peut donc s'écrire de façon unique (à l'ordre près) sous la forme :

$$P = \lambda \prod_{i=1}^{r} (X - a_i)^{\alpha_i}$$

où $\lambda \in \mathbb{C}^*$, les a_i sont des complexes distincts deux à deux et les α_i des entiers naturels non nuls.

. **Décomposition en produit de polynômes irréductibles dans $\mathbb{R}[X]$**
Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 à discriminant strictement négatif.
Tout polynôme P non constant de $\mathbb{R}[X]$ peut donc s'écrire de façon unique (à l'ordre près) sous la forme :

$$P = \lambda \prod_{i=1}^{r} (X - a_i)^{\alpha_i} \prod_{j=1}^{s} (X^2 + b_j X + c_j)^{\beta_j}$$

où $\lambda \in \mathbb{R}^*$, les a_i sont des réels distincts deux à deux, les $X^2 + b_j X + c_j$ des trinômes distincts deux à deux et à discriminant strictement négatif, et les α_i, β_j des entiers naturels non nuls.

[S1.14] Corps

• Un corps est un anneau commutatif, non réduit à $\{0\}$ dans lequel tout élément non nul est inversible.
Autrement dit, un triplet $(\mathbb{K}, +, \times)$ est un corps si et seulement si :

(i) $(\mathbb{K},+)$ est un groupe abélien ;

(ii) $(\mathbb{K}^*,\times)$ est un groupe abélien (où $\mathbb{K}^* = \mathbb{K} \setminus \{0_K\}$) ;

(iii) $\times$ est distributive par rapport à $+$.

- Si $(\mathbb{K}, +, \times)$ est un corps, l'anneau $(\mathbb{K}, +, \times)$ est intègre.
- Un anneau commutatif $(A, +, \times)$ non réduit à $\{0\}$ est un corps si et seulement si ses seuls idéaux sont $\{0\}$ et A.

  ✓ Ce résultat montre que la notion d'idéal n'a pas d'intérêt dans un corps !
- Tout anneau *intègre* et *fini* (non réduit à $\{0\}$) est un corps.
- L'anneau $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ est un corps si et seulement si p est un nombre premier. Dans ce cas, $\mathcal{U}(\mathbb{Z}/p\mathbb{Z}) = \mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}$ est un groupe multiplicatif d'ordre $p-1$.
  On en déduit le **petit théorème de Fermat** :

  si p est premier et si $a \in \mathbb{Z}$ n'est pas multiple de p, $a^{p-1} \equiv 1 \pmod{p}$.

  Plus généralement, si p est premier, on a :

  $$\forall\, a \in \mathbb{Z},\ a^p \equiv a \pmod{p}\,.$$

- **Sous-corps**

  Soient $(\mathbb{K}, +, \times)$ un corps, et $\mathbb{L}$ une partie de $\mathbb{K}$. On dit que $\mathbb{L}$ est un sous-corps de $\mathbb{K}$ si $\mathbb{L}$ est un sous-anneau de $\mathbb{K}$ qui, pour les lois induites, est aussi un corps.
  Pour qu'une partie $\mathbb{L}$ d'un corps $(\mathbb{K}, +, \times)$ soit un sous-corps de $\mathbb{K}$, il faut et il suffit que :

  (i) $1_{\mathbb{K}} \in \mathbb{L}$;

  (ii) $\forall\,(x,y) \in \mathbb{L}^2,\ x-y \in \mathbb{L}$ et $xy \in \mathbb{L}$;

  (iii) $\forall\, x \in \mathbb{L} \setminus \{0_{\mathbb{K}}\},\ x^{-1} \in \mathbb{L}$.

  Exemple : $\mathbb{Q} + \sqrt{3}\mathbb{Q} = \{a + b\sqrt{3}\ ;\ (a,b) \in \mathbb{Q}^2\}$ est un sous-corps de $\mathbb{R}$.

- **Morphismes de corps**

  Soient $\mathbb{K}$ et $\mathbb{K}'$ deux corps (dont les lois sont notées de façon identique).
  Une application f de $\mathbb{K}$ dans $\mathbb{K}'$ est appelée un morphisme de corps si f est un morphisme pour les structures d'anneaux sous-jacentes, c'est-à-dire :

  (i) $f(1_{\mathbb{K}}) = 1_{\mathbb{K}'}$;

  (ii) $\forall\,(x,y) \in \mathbb{K}^2,\ f(x+y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$.

  Si $f : \mathbb{K} \to \mathbb{K}'$ est un morphisme de corps, alors f est injectif.

[S1.15] Algèbres

- Soit $(\mathbb{K}, +, \times)$ un corps.
  Une algèbre sur $\mathbb{K}$ est un *quadruplet* $(E, +, ., \times)$, où E est un ensemble muni de deux lois de composition internes $+$ et $\times$ et d'une loi de composition externe $.$ de domaine d'opérateurs $\mathbb{K}$, telles que :

(i) $(E, +, \times)$ est un anneau ;

(ii) $(E, +, .)$ est un $\mathbb{K}$-espace vectoriel (les espaces vectoriels seront étudiés en détail dans le thème 2) ;

(iii) $\forall \lambda \in \mathbb{K}, \forall (x,y) \in E^2, \lambda.(x \times y) = (\lambda.x) \times y = x \times (\lambda.y)$.

L'algèbre $(E,+,.,\times)$ est dite commutative lorsque l'anneau $(E,+,\times)$ l'est, c'est-à-dire si la loi $\times$ est commutative.

Exemples

- L'ensemble des polynômes $\mathbb{K}[X]$ à coefficients dans $\mathbb{K}$ est une $\mathbb{K}$-algèbre commutative et intègre (pour les lois habituelles, rappelées en **[S1.12]**).
- Si E est un $\mathbb{K}$-espace vectoriel, l'ensemble $(\mathscr{L}(E),+,\cdot,\circ)$ des endomorphismes de E est une $\mathbb{K}$-algèbre, non commutative, non intègre, dès que $\dim E \geqslant 2$.
- Si $n \in \mathbb{N}^*$, l'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$ est une $\mathbb{K}$-algèbre pour les lois usuelles sur les matrices.

- **Sous-algèbre**

Soit E une $\mathbb{K}$-algèbre. Une partie F de E est une sous-algèbre de E si $(F,+,.,\times)$ est encore une $\mathbb{K}$-algèbre.

Cela équivaut à dire que F est à la fois un sous-anneau de $(E, +, \times)$ et un sous-espace vectoriel de $(E, +, .)$

Caractérisation d'une sous-algèbre

Soit E une $\mathbb{K}$-algèbre. Une partie F de E est une sous-algèbre de E si et seulement si :

(i) $1_E \in F$;

(ii) $\forall (x,y) \in F^2, x + y \in F$ et $xy \in F$;

(iii) $\forall \lambda \in \mathbb{K}, \forall x \in F, \lambda.x \in F$.

- **Morphisme d'algèbres**

▪ Soit $(\mathbb{K}, +, \times)$ un corps et E et E' deux $\mathbb{K}$-algèbres (dont les lois sont notées de façon identique).

Une application f de E dans E' est appelée un morphisme d'algèbres si :

(i) $f(1_E) = 1_{E'}$;

(ii) $\forall (x,y) \in E^2, f(x + y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$;

(iii) $\forall \lambda \in \mathbb{K}, \forall x \in E\ , f(\lambda.x) = \lambda.f(x)$.

f est donc à la fois un morphisme d'espaces vectoriels (c'est-à-dire une application linéaire), et un morphisme d'anneaux.

▪ Soit $f : A \to A'$ un morphisme de $\mathbb{K}$-algèbres. Alors :

- si B est une sous-algèbre de A, $f(B)$ est une sous-algèbre de A' ;
- si B' est une sous-algèbre de A', son image réciproque $f^{-1}(B')$ est une sous-algèbre de A.

Thème 2 - Espaces vectoriels et applications linéaires

Dans tout le chapitre, $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$, le plus souvent $\mathbb{R}$ ou $\mathbb{C}$.

[S2.1] Définition d'un espace vectoriel

Un $\mathbb{K}$-espace vectoriel est un triplet $(E, +, .)$, où E est un ensemble muni :

- d'une loi de composition interne sur E, notée $+$, telle que $(E,+)$ soit un groupe abélien ;
- d'une loi de composition externe sur $\mathbb{K}$, notée $.$, c'est-à-dire une application $\begin{cases} \mathbb{K} \times E & \longrightarrow & E \\ (\lambda,x) & \longmapsto & \lambda.x \end{cases}$ vérifiant les propriétés suivantes :

$$\forall\,(\lambda,\mu) \in \mathbb{K}^2,\ \forall\,(x,y) \in E^2,\ \begin{cases} \text{(i)} & 1_{\mathbb{K}}.x = x\,; \\ \text{(ii)} & \lambda.(x+y) = \lambda.x + \lambda.y\,; \\ \text{(iii)} & (\lambda+\mu).x = \lambda.x + \mu.x\,; \\ \text{(iv)} & \lambda.(\mu.x) = (\lambda\mu).x\,. \end{cases}$$

Exemples

Il est important de connaître un certain nombre d'espaces vectoriels classiques.

- Si D est un ensemble quelconque et E un $\mathbb{K}$-espace vectoriel, l'ensemble $\mathcal{A}(D,E)$ des applications de D dans E peut être muni d'une structure de $\mathbb{K}$-espace vectoriel pour les lois $+$ et $.$ définies par :

  $$(f,g) \mapsto f+g \text{ avec : } \forall\, x \in D,\ (f+g)(x) = f(x) + g(x)\,;$$
  $$(\lambda,f) \mapsto \lambda.f \text{ avec : } \forall\, x \in D,\ (\lambda.f)(x) = \lambda.f(x)\,.$$

  En particulier, l'ensemble $E^{\mathbb{N}}$ des suites à valeurs dans E est un $\mathbb{K}$-espace vectoriel.
- $\mathbb{K}[X]$, ensemble des polynômes à coefficients dans $\mathbb{K}$, est un $\mathbb{K}$-espace vectoriel.
- $\mathcal{M}_{p,q}(\mathbb{K})$, ensemble des matrices à p lignes et q colonnes à coefficients dans $\mathbb{K}$, est un $\mathbb{K}$-espace vectoriel.

[S2.2] Espace vectoriel produit

Soient $E_1, E_2, \ldots, E_p$ p $\mathbb{K}$-espaces vectoriels.
L'ensemble produit $E = E_1 \times E_2 \times \ldots \times E_p$ est un $\mathbb{K}$-espace vectoriel pour les lois définies par :

$$\begin{aligned} (x_1,x_2,\ldots,x_p) + (y_1,y_2,\ldots,y_p) &= (x_1+y_1, x_2+y_2, \ldots, x_p+y_p)\,; \\ \lambda.(x_1,x_2,\ldots,x_p) &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_p)\,. \end{aligned}$$

Muni de ces lois, E s'appelle l'espace vectoriel produit des E_i. Le vecteur nul de E est le p-uplet $(0_{E_1},\ldots,0_{E_p})$.

Exemple : $\mathbb{K}^p$ est un $\mathbb{K}$-espace vectoriel pour les lois définies ci-dessus.

[S2.3] Sous-espace vectoriel

- Soit $(E, +, .)$ un $\mathbb{K}$-espace vectoriel. Par définition, une *partie* F de E est un sous-espace vectoriel de E si $(F, +, .)$ est encore un $\mathbb{K}$-espace vectoriel.
- Cela équivaut à dire que F est une partie non vide de E stable par combinaisons linéaires, soit :

$$F \subset E, \quad F \neq \emptyset \quad \text{et} \quad \forall \lambda \in \mathbb{K}, \forall (x,y) \in F^2 \ : \ x+y \in F \text{ et } \lambda x \in F$$

ou encore :

$$F \subset E, \quad F \neq \emptyset \quad \text{et} \quad \forall \lambda \in \mathbb{K}, \forall (x,y) \in F^2 \ : \ \lambda x + y \ \in \ F\,.$$

- **Exemples**

  Il existe un certain nombre de sous-espaces vectoriels classiques (donc d'espaces vectoriels) à connaître.

  - $\{0_E\}$ et E sont des sous-espaces vectoriels de E (dits triviaux).
  - Si $n \in \mathbb{N}$, $\mathbb{K}_n[X]$, ensemble des polynômes de $\mathbb{K}[X]$ de degré inférieur ou égal à n, est un sous-espace vectoriel de $\mathbb{K}[X]$.

    ✓ Attention : l'ensemble des polynômes de degré *exactement* n n'est PAS un sous-espace vectoriel de $\mathbb{K}[X]$; en effet, il ne contient même pas le polynôme nul !
  - Si $P \in \mathbb{K}[X]$, l'ensemble $\mathbb{K}[X]\cdot P$ des multiples de P est un sous-espace vectoriel de $\mathbb{K}[X]$ (c'est l'idéal engendré par P).
  - Si I est un intervalle de $\mathbb{R}$, l'ensemble $\mathscr{C}(I,\mathbb{R})$ des fonctions continues sur I à valeurs réelles est un sous-espace vectoriel de $\mathcal{A}(I,\mathbb{R})$.

[S2.4] Intersection de sous-espaces vectoriels

Si $(E_i)_{i\in I}$ est une famille quelconque de sous-espaces vectoriels de E, leur intersection $\bigcap_{i\in I} E_i$ est encore un sous-espace vectoriel de E.

✓ La *réunion* de sous-espaces vectoriels de E n'est pas, en général, un sous-espace vectoriel de E.
Plus précisément, on peut démontrer que si A et B sont deux sous-espaces vectoriels d'un espace vectoriel E, $A \cup B$ est encore un sous-espace vectoriel de E si et seulement si $A \subset B$ ou $B \subset A$.

[S2.5] Sous-espace vectoriel engendré, combinaisons linéaires

- Soit X une partie quelconque d'un $\mathbb{K}$-espace vectoriel E.
  L'intersection de tous les sous-espaces vectoriels de E qui contiennent X est encore un sous-espace vectoriel de E ; c'est en fait le plus petit sous-espace vectoriel de E contenant X (au sens de l'inclusion) ; on l'appelle sous-espace vectoriel engendré par X, et on le note $\mathrm{Vect}(X)$.

- Si $X = \emptyset$, $\text{Vect}(\emptyset) = \{0_E\}$. Sinon, $\text{Vect}(X)$ est exactement l'ensemble des combinaisons linéaires des éléments de X, c'est-à-dire les éléments de la forme :

$$\sum_{i=1}^{n} \lambda_i x_i \quad \text{avec } n \in \mathbb{N}^*,\ \lambda_i \in \mathbb{K} \text{ et } x_i \in X \text{ pour tout } i\,.$$

  ✓ Par définition, une combinaison linéaire est toujours une somme *finie*.

  ✓ Le résultat précédent montre, en particulier, que l'ensemble des combinaisons linéaires des vecteurs d'une partie de E est un sous-espace vectoriel de E. Cela peut être utile pour abréger certaines démonstrations.

[S2.6] Famille génératrice

Une famille $(x_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est dite génératrice de E si le sous-espace vectoriel qu'elle engendre est égal à l'espace E tout entier :

$$E = \text{Vect}\big((x_i)_{i\in I}\big)\,,$$

c'est-à-dire si tout vecteur de E est combinaison linéaire des x_i .

Toute famille contenant une famille génératrice de E est encore génératrice de E.

[S2.7] Dépendance et indépendance linéaire

- Une famille finie $(x_i)_{1\leqslant i\leqslant n}$ d'un $\mathbb{K}$-espace vectoriel E est dite libre s'il n'existe pas de relation de dépendance linéaire non triviale entre ces vecteurs, c'est-à-dire si, pour toute famille $(\lambda_i)_{1\leqslant i\leqslant n}$ d'éléments de $\mathbb{K}$:

$$\sum_{i=1}^{n} \lambda_i x_i = 0_E \implies \forall\, i \in [\![1\,;n]\!],\ \lambda_i = 0_{\mathbb{K}}$$

  (on dit aussi que les x_i sont linéairement indépendants).

- Une famille quelconque $(x_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est dite libre si *toutes* ses sous-familles finies sont libres.

  Exemple : dans $\mathbb{K}[X]$, toute famille de polynômes non nuls de degrés distincts est libre.

- Une famille $(x_i)_{i\in I}$ d'un $\mathbb{K}$-espace vectoriel E est dite liée si elle n'est pas libre, c'est-à-dire s'il existe une famille $(\lambda_i)_{i\in I}$ d'éléments de $\mathbb{K}$, à support fini (c'est-à-dire tous nuls sauf un nombre fini), *non tous nuls*, telle que $\sum_{i\in I} \lambda_i x_i = 0_E$ (on dit aussi que les x_i sont linéairement dépendants).

- **Propriétés**
  - Une famille réduite à un élément $\{x\}$ est libre si et seulement si $x \neq 0_E$.
  - Toute famille contenant 0_E est liée.
  - Toute sous-famille d'une famille libre est libre. Toute famille contenant une famille liée est liée.
  - Une famille $(x_i)_{i\in I}$ est liée si et seulement si il existe $j \in I$ tel que x_j soit combinaison linéaire des $(x_i)_{i\in I, i\neq j}$.
    Une famille $(x_i)_{i\in I}$ est libre si aucun des x_i n'est combinaison linéaire des autres.

[S2.8] Bases

Une famille $\mathscr{B}$ d'un $\mathbb{K}$-espace vectoriel E s'appelle une base de E si elle est à la fois libre et génératrice de E.

Cela équivaut à dire que tout vecteur x de E s'écrit de manière unique comme combinaison linéaire des éléments de $\mathscr{B}$. Les coefficients de cette combinaison linéaire sont les coordonnées de x dans la base $\mathscr{B}$.

[S2.9] Somme de sous espaces vectoriels

Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de n sous-espaces vectoriels de E ($n \in \mathbb{N}^*$).

On appelle somme des E_i, et on note $\sum_{i=1}^{n} E_i$, l'ensemble des vecteurs de la forme $\sum_{i=1}^{n} x_i$, où $x_i \in E_i$ pour tout $i \in [\![1\,;n]\!]$.

C'est aussi le sous-espace vectoriel de E engendré par la réunion des E_i, c'est-à-dire le plus petit sous-espace vectoriel de E qui contient tous les E_i :

$$\sum_{i=1}^{n} E_i = \mathrm{Vect}\left(\bigcup_{1\leqslant i\leqslant n} E_i \right).$$

✓ On prendra soin de ne pas confondre la *réunion*, qui est une notion ensembliste, et la *somme* de sous-espaces vectoriels.

Exemple : si $x_1,\ldots,x_n$ sont n vecteurs de E, $\sum_{i=1}^{n} \mathbb{K}.x_i = \mathrm{Vect}(x_1,\ldots,x_n)$.

[S2.10] Somme directe de sous-espaces vectoriels

- Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de n sous-espaces vectoriels de E.

  On dit que la somme $\sum_{i=1}^{n} E_i$ est directe lorsque pour tout $x \in \sum_{i=1}^{n} E_i$ l'écriture $x = \sum_{i=1}^{n} x_i$ avec $x_i \in E_i$ pour tout i est *unique*. On note alors : $\sum_{i=1}^{n} E_i = \bigoplus_{i=1}^{n} E_i$.

  Exemple : si $x_1,\ldots,x_n$ sont n vecteurs *non nuls* de E, dire que la somme $\sum_{i=1}^{n} \mathbb{K}.x_i$ est directe équivaut à dire que la famille $(x_1,\ldots,x_n)$ est libre.

- **Caractérisations d'une somme directe**

  . Pour que la somme $\sum_{1\leqslant i\leqslant n} E_i$ soit directe, il faut et il suffit que pour toute famille $(x_i)_{1\leqslant i\leqslant n}$ avec $x_i \in E_i$ pour tout i on ait :

  $$\sum_{i=1}^{n} x_i = 0 \implies \forall\, i \in [\![1\,;n]\!],\ x_i = 0\,.$$

  ✓ Il s'agit de la caractérisation la plus simple à utiliser pour montrer que n sous-espaces vectoriels sont en somme directe dès que $n \geqslant 3$.

- Pour que la somme $\sum_{i=1}^{n} E_i$ soit directe, il faut et il suffit qu'il existe $j \in [\![1\,;n]\!]$ tel que $\sum_{\substack{1\leqslant i\leqslant n\\ i\neq j}} E_i$ soit directe et que les sous-espaces $\sum_{\substack{1\leqslant i\leqslant n\\ i\neq j}} E_i$ et E_j soient en somme directe (et dans ce cas, cette propriété est vraie pour tout $j \in [\![1\,;n]\!]$).

  ✓ Cette caractérisation est moins utilisée, mais elle peut être pratique dans des démonstrations par récurrence.

- **Somme directe de deux sous-espaces**

  Dans le cas de *deux* sous-espaces vectoriels (et dans ce cas seulement), on a une caractérisation plus simple : pour que la somme de deux sous-espaces vectoriels E_1 et E_2 de E soit directe, il faut et il suffit que $E_1 \cap E_2 = \{0_E\}$.

- **Sous-espaces supplémentaires**

  Les sous-espaces vectoriels E_i $(1 \leqslant i \leqslant n)$ de E sont dits supplémentaires si $E = \bigoplus_{i=1}^{n} E_i$. Cela équivaut à dire que tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = \sum_{i=1}^{n} x_i$ avec $x_i \in E_i$ pour tout i.

  Exemple : si $x_1, \ldots, x_n$ sont n vecteurs *non nuls* de E, dire que $E = \bigoplus_{i=1}^{n} \mathbb{K}.x_i$ équivaut à dire que la famille $(x_1, \ldots, x_n)$ est une base de E.

- **Base adaptée à une décomposition en somme directe**

  - Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{i=1}^{n} E_i$.

    Si, pour tout $i \in [\![1\,;n]\!]$, $\mathscr{B}_i$ est une base de E_i, alors les $\mathscr{B}_i$ sont deux à deux disjointes et $\mathscr{B} = \bigcup_{i=1}^{n} \mathscr{B}_i$ est une base de E.

  - Soit $\mathscr{B}$ une base de E, et $(\mathscr{B}_i)_{1\leqslant i\leqslant n}$ une partition de $\mathscr{B}$.

    Si, pour tout $i \in [\![1\,;n]\!]$, $E_i = \mathrm{Vect}(\mathscr{B}_i)$, alors $E = \bigoplus_{i=1}^{n} E_i$.

[S2.11] Espaces vectoriels de dimension finie

- On dit qu'un $\mathbb{K}$-espace vectoriel est de dimension finie s'il possède une famille génératrice finie.

- **Théorème de la base incomplète**

  Soient L une famille libre et G une famille génératrice de E, espace vectoriel de dimension finie.

  Alors il existe une base $\mathscr{B}$ de E telle que : $L \subset \mathscr{B} \subset L \cup G$ (*autrement dit, on peut compléter L à l'aide de vecteurs de G pour obtenir une base*).

  On en déduit que tout espace vectoriel de dimension finie E possède (au moins) une base (dans le cas où $E = \{0\}$, on peut convenir qu'une base de E est $\emptyset$).

- **Dimension**

  Toutes les bases d'un espace vectoriel E de dimension finie ont le même nombre d'éléments, appelé la dimension de E, notée $\dim_{\mathbb{K}} E$ ou plus simplement $\dim E$.

- **Caractérisation des bases**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
  - Toute famille libre de E a au plus n éléments ; et si elle en a exactement n, c'est une base de E.
  - Toute famille génératrice de E a au moins n éléments ; et si elle en a exactement n, c'est une base de E.

  Exemple : toute famille de polynômes de degrés échelonnés de 0 à n est une base de $\mathbb{K}_n[X]$.
- **Dimension d'un espace vectoriel produit**
  Soient $E_1, E_2, \ldots, E_p$ p $\mathbb{K}$-espaces vectoriels de dimensions finies.
  Alors l'espace vectoriel produit $E = E_1 \times E_2 \times \ldots \times E_p$ est de dimension finie et $\dim E = \sum_{i=1}^{p} \dim(E_i)$.

[S2.12] Dimension d'un sous-espace vectoriel

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soit F un sous-espace vectoriel de E. Alors :
  - F est de dimension finie, et $\dim F \leqslant \dim E$;
  - de plus, si $\dim E = \dim F$, alors $F = E$.
- **Rang d'une famille**
  Soit $(x_i)_{i\in I}$ une famille d'éléments d'un $\mathbb{K}$-espace vectoriel E.
  Si le sous-espace vectoriel qu'elle engendre est de dimension finie, on définit le rang de la famille $(x_i)_{i\in I}$ par :
  $$\mathrm{rg}\big((x_i)_{i\in I}\big) = \dim\big(\mathrm{Vect}((x_i)_{i\in I})\big).$$
- **Propriétés**
  - Si E est de dimension finie, alors : $\mathrm{rg}\big((x_i)_{i\in I}\big) \leqslant \dim(E)$.
  - Si $(x_1,x_2,\ldots,x_p)$ est une famille finie, alors : $\mathrm{rg}(x_1,x_2,\ldots,x_p) \leqslant p$.

[S2.13] Dimension de la somme de deux sous-espaces

- **Dimension d'un supplémentaire**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
  - Tout sous-espace vectoriel de E possède (au moins) un supplémentaire.
  - Si $E = F \oplus G$, $\dim F + \dim G = \dim E$.
- **Formule de Grassmann**
  Soit E un $\mathbb{K}$-espace vectoriel, et soient F et G deux sous-espaces vectoriels de dimensions finies de E. Alors $F + G$ est de dimension finie et :
  $$\dim(F+G) = \dim F + \dim G - \dim(F \cap G).$$

- **Caractérisation de deux sous-espaces vectoriels supplémentaires**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et soient F et G deux sous-espaces vectoriels de E.
  Pour que F et G soient supplémentaires, il faut et il suffit que :

$$F + G = E \quad \text{et} \quad \dim E = \dim F + \dim G$$

  ou bien que :

$$F \cap G = \{0_E\} \quad \text{et} \quad \dim E = \dim F + \dim G\,.$$

[S2.14] Dimension de la somme de n sous-espaces

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $(E_i)_{1\leqslant i\leqslant n}$ une famille de sous-espaces vectoriels de E. On a l'inégalité :

$$\dim\left(\sum_{i=1}^{n} E_i\right) \leqslant \sum_{i=1}^{n} \dim E_i\,,$$

  et il y a égalité si et seulement si la somme est directe.
- **Caractérisation de sous-espaces vectoriels supplémentaires**
  Pour que les E_i soient supplémentaires, il faut et il suffit que :

$$\text{la somme } \sum_{i=1}^{n} E_i \text{ soit directe et que } \dim E = \sum_{i=1}^{n} \dim(E_i)$$

  ou bien que :

$$E = \sum_{i=1}^{n} E_i \quad \text{et} \quad \dim E = \sum_{i=1}^{n} \dim(E_i)\,.$$

[S2.15] Applications linéaires

- Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $u\colon E \to F$ est dite linéaire si :

$$\forall\,\lambda \in \mathbb{K},\ \forall\,(x,y) \in E^2,\ u(x+y) = u(x) + u(y) \quad \text{et} \quad u(\lambda x) = \lambda u(x)\,.$$

  On peut remplacer cette définition par :

$$\forall\,\lambda \in \mathbb{K},\ \forall\,(x,y) \in E^2,\ u(\lambda x + y) = \lambda u(x) + u(y)\,.$$

  Si u est linéaire, on aura plus généralement, pour toute famille de scalaires $(\lambda_i)_{1\leqslant i\leqslant n}$ et toute famille $(x_i)_{1\leqslant i\leqslant n}$ de vecteurs de E :

$$u\left(\sum_{i=1}^{n} \lambda_i x_i\right) = \sum_{i=1}^{n} \lambda_i u(x_i)\,.$$

- Un isomorphisme de E sur F est une application linéaire bijective de E sur F ; un endomorphisme de E est une application linéaire de E dans lui-même ; un automorphisme de E est un endomorphisme de E bijectif.

- On note $\mathscr{L}(E,F)$ l'ensemble des applications linéaires de E dans F, $\mathscr{L}(E)$ l'ensemble des endomorphismes de E et $\mathrm{GL}(E)$ celui des automorphismes de E.
- **Exemples**
  De même qu'il faut connaître certains espaces vectoriels de référence, il faut aussi connaître certaines applications linéaires classiques. Citons-en quelques-unes.
  - L'application nulle de E dans F, qui à tout $x \in E$ associe 0_F.
  - L'application identique de E : $\forall x \in E$, $\mathrm{Id}_E(x) = x$.
  - L'homothétie de rapport $\lambda \in \mathbb{K}$: $h_\lambda : x \mapsto \lambda x \quad (h_\lambda = \lambda.\mathrm{Id}_E)$.
  - L'application $P \mapsto P'$ de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$.
  - L'application $f \mapsto f(a)$ de $\mathcal{A}(D,E)$ dans E, avec $a \in D$.
  - L'application $f \mapsto \displaystyle\int_a^b f(t)\,\mathrm{d}t$ de $\mathscr{C}([a\,;b],\mathbb{R})$ dans $\mathbb{R}$.

[S2.16] Opérations sur les applications linéaires

- Soient E et F deux $\mathbb{K}$-espaces vectoriels. Alors $\mathscr{L}(E,F)$ est un sous-espace vectoriel de $\mathcal{A}(E,F)$. Autrement dit, si u et v sont deux applications linéaires de E dans F et si λ est un scalaire, l'application $\lambda u + v$ est linéaire.
  De plus, lorsque E et F sont de dimensions finies, on a :
  $$\dim \mathscr{L}(E,F) = \dim E \times \dim F\,.$$
- Soient E, F et G trois $\mathbb{K}$-espaces vectoriels.
  Si $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$, alors $v \circ u \in \mathscr{L}(E,G)$.
- Si u est un isomorphisme de E sur F, l'application réciproque u^{-1} est encore linéaire (ce sera donc un isomorphisme de F sur E).
- $\big(\mathscr{L}(E),\ +\ ,.\ ,\ \circ\ \big)$ est une algèbre (non commutative, non intègre, dès que $\dim E \geqslant 2$).
  $\big(\mathrm{GL}(E), \circ\big)$ est un groupe ; c'est le groupe des éléments inversibles de l'anneau $\big(\mathscr{L}(E), +\,, \circ\big)$.

[S2.17] Image directe et image réciproque d'un sous-espace, noyau

- Soient E et F deux $\mathbb{K}$-espaces vectoriels, et $u \in \mathscr{L}(E,F)$.
  - Si E' est un sous-espace vectoriel de E, son image $u(E')$ par u est un sous-espace vectoriel de F.
  - Si F' est un sous-espace vectoriel de F, son image réciproque par u :
    $$u^{-1}(F') = \{x \in E \mid u(x) \in F'\}\,,$$
    est un sous-espace vectoriel de E.
- Soit $u \in \mathscr{L}(E,F)$. L'ensemble image de E par u est un sous-espace vectoriel de F, appelé image de u, et noté $\mathrm{Im}\,u$. Ainsi :
  $$\mathrm{Im}\,u = \{u(x) \mid x \in E\}\,.$$

- Soit $u \in \mathscr{L}(E,F)$. L'image réciproque de $\{0_F\}$ par u est un sous-espace vectoriel de E, appelé noyau de u, et noté $\operatorname{Ker} u$. Ainsi :

$$\operatorname{Ker} u = u^{-1}(\{0_F\}) = \{x \in E \mid u(x) = 0_F\}\,.$$

- **Théorème**
  Soit $u \in \mathscr{L}(E,F)$. Alors :

$$u \text{ surjective} \iff \operatorname{Im} u = F \quad ; \quad u \text{ injective} \iff \operatorname{Ker} u = \{0_E\}\,.$$

- **Deux propriétés simples mais importantes**
  Soient E, F et G trois $\mathbb{K}$-espaces vectoriels, $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$. Alors :
  - $\operatorname{Im}(v \circ u) \subset \operatorname{Im} v$ et $\operatorname{Ker}(v \circ u) \supset \operatorname{Ker} u$.
  - $v \circ u = 0 \iff \operatorname{Im} u \subset \operatorname{Ker} v$.

- **Noyau d'une restriction**
  Si $u \in \mathscr{L}(E,F)$ et si E_1 est un sous-espace vectoriel de E, la restriction de u à E_1 a pour noyau :

$$\operatorname{Ker}(u|_{E_1}) = \operatorname{Ker} u \cap E_1\,.$$

- **Théorème d'isomorphisme**
  Soit $u \in \mathscr{L}(E,F)$. La restriction de u à tout supplémentaire de $\operatorname{Ker} u$ dans E est un isomorphisme de ce supplémentaire sur $\operatorname{Im} u$.

[S2.18] Détermination d'une application linéaire

- Soit $(E_i)_{1\leqslant i\leqslant n}$ une famille de sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E, telle que $E = \bigoplus_{i=1}^{n} E_i$, et soit F un $\mathbb{K}$-espace vectoriel.
  Pour tout $i \in [\![1\,;n]\!]$, soit u_i une application linéaire de E_i dans F.
  Alors il existe une et une seule application linéaire $u \in \mathscr{L}(E,F)$ telle que pour tout $i \in [\![1\,;n]\!]$, $u_i = u|_{E_i}$.
  (*autrement dit, une application linéaire est entièrement déterminée par ses restrictions à des sous-espaces vectoriels supplémentaires*).

- Soient E et F deux $\mathbb{K}$-espaces vectoriels, $(e_i)_{i\in I}$ une base de E et $(b_i)_{i\in I}$ une famille de vecteurs de F.
  Alors il existe une et une seule application linéaire $u \in \mathscr{L}(E,F)$ telle que $u(e_i) = b_i$ pour tout $i \in I$.
  (*autrement dit, une application linéaire est entièrement déterminée par les images des vecteurs d'une base*).

- Soient E et F deux $\mathbb{K}$-espaces vectoriels et $u \in \mathscr{L}(E,F)$.
  u est injective (respectivement surjective, respectivement bijective) si et seulement si l'image d'*une* base par u est une famille libre dans F (respectivement génératrice de F, respectivement est une base de F) et, lorsque cela est le cas, cela vaut pour *toutes* les bases.

- Il est important de retenir que, si $u \in \mathscr{L}(E,F)$ et si $(e_i)_{i\in I}$ est une base de E, alors $\operatorname{Im} u$ est le sous-espace vectoriel de F engendré par les vecteurs $u(e_i)$.

[S2.19] Rang d'une application linéaire

- Soient E et F deux $\mathbb{K}$-espaces vectoriels et $u \in \mathscr{L}(E,F)$. Si $\operatorname{Im} u$ est de dimension finie, sa dimension s'appelle le rang de u, noté $\operatorname{rg} u$.
  Si $(e_i)_{i\in I}$ est une base de E, le rang de u est aussi celui de la famille $\big(u(e_i)\big)_{i\in I}$.
- **Théorème du rang**
  Soient E et F deux $\mathbb{K}$-espaces vectoriels, avec E de dimension finie, et soit $u \in \mathscr{L}(E,F)$. Alors $\operatorname{Im} u$ est de dimension finie et :
  $$\dim(\operatorname{Im} u) + \dim(\operatorname{Ker} u) \;=\; \dim E\,.$$
- **Conséquences**
  u désigne ici une application linéaire de E dans F.
  - Si E est de dimension finie :
    $$\operatorname{rg} u \leqslant \dim E \quad \text{et} \quad \operatorname{rg} u = \dim E \iff u \text{ injective}\,.$$
  - Si F est de dimension finie :
    $$\operatorname{rg} u \leqslant \dim F \quad \text{et} \quad \operatorname{rg} u = \dim F \iff u \text{ surjective}\,.$$
  - Si E et F sont de dimensions finies et **si dimE = dimF**, on a :
    $$u \text{ injective} \iff u \text{ surjective} \iff u \text{ bijective.}$$
  - Soient E et F de dimensions finies, avec $\dim E = \dim F$. Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,E)$ tels que $v \circ u = \mathrm{Id}_E$ (ou $u \circ v = \mathrm{Id}_F$). Alors u et v sont des isomorphismes, réciproques l'un de l'autre.

    ✓ Attention : les deux dernières propriétés ci-dessus ne sont plus vraies lorsque E et F ne sont pas de *même* dimension, ou bien lorsque l'un d'eux n'est pas de dimension finie.
- **Rang de la composée**
  Soient E, F et G trois $\mathbb{K}$-espaces vectoriels, E et F étant de dimensions finies.
  Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$. Alors on a :
  - $\operatorname{rg}(v \circ u) \leqslant \operatorname{rg} v$ et, si u est surjective, il y a égalité.
  - $\operatorname{rg}(v \circ u) \leqslant \operatorname{rg} u$ et, si v est injective, il y a égalité.

  En particulier, il y a invariance du rang par la composition avec un isomorphisme.

[S2.20] Formes linéaires, hyperplans

- **Espace dual**
  - On rappelle qu'une forme linéaire sur un $\mathbb{K}$-espace vectoriel E est une application linéaire de E dans $\mathbb{K}$.
    L'espace vectoriel $\mathscr{L}(E,\mathbb{K})$ des formes linéaires sur E s'appelle l'espace vectoriel dual de E, et est noté E^*.

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, et $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E.

  Pour tout $i \in [\![1\,;n]\!]$, l'application e_i^* qui à tout $x = \sum_{i=1}^{n} x_i e_i$ de E, associe $e_i^*(x) = x_i$, est une forme linéaire sur E, appelée i^{e} forme linéaire coordonnée dans $\mathscr{B}$.

  e_i^* est l'unique forme linéaire sur E telle que :

$$\forall\, j \in [\![1\,;n]\!],\ e_i^*(e_j) = \delta_{ij}\,.$$

- La famille $\mathscr{B}^* = (e_1^*, \ldots, e_n^*)$ est une base de E^*, appelée base duale de $\mathscr{B}$.

- **Hyperplan**

  Soit E un $\mathbb{K}$-espace vectoriel. Un hyperplan H de E est un sous-espace vectoriel de E qui admet pour supplémentaire une droite vectorielle D : $E = D \oplus H$.

  Si H est un hyperplan, alors, pour toute droite vectorielle D' non incluse dans H, on a $E = D' \oplus H$.

  Lorsque E est de dimension finie $n \geqslant 1$, un sous-espace vectoriel H de E est un hyperplan si et seulement si $\dim H = n - 1$.

- **Lien avec les formes linéaires**

  - Un sous-espace vectoriel H de E est un hyperplan si et seulement si il existe une forme linéaire φ sur E, non nulle, telle que $H = \operatorname{Ker}\varphi$.
  - Si φ et ψ sont deux formes linéaires non nulles sur E telles que $\operatorname{Ker}\varphi = \operatorname{Ker}\psi$, alors il existe $\lambda \in \mathbb{K}$ tel que $\psi = \lambda\varphi$.

  Pour tout hyperplan H, il existe donc une forme linéaire non nulle φ telle que $H = \{x \in E\,, \varphi(x) = 0\}$. La relation $\varphi(x) = 0$ s'appelle une équation de H.

- **Expression analytique**

  Lorsque E est de dimension finie n, muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, toute forme linéaire φ sur E est de la forme :

$$\varphi = \sum_{i=1}^{n} a_i e_i^*\,, \text{ avec } a_i = \varphi(e_i) \in \mathbb{K}\,.$$

  Cela équivaut à dire que pour tout $x = \sum_{i=1}^{n} x_i e_i \in E$, $\varphi(x) = \sum_{i=1}^{n} a_i x_i$.

  Il en résulte que tout hyperplan possède une équation de la forme :

$$\sum_{i=1}^{n} a_i x_i = 0 \quad (a_i \text{ scalaires non tous nuls})\,.$$

- **Intersection d'hyperplans**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

  Si $(\varphi_1, \ldots, \varphi_p)$ est une famille de p formes linéaires *linéairement indépendantes* sur E ($p \in \mathbb{N}^*$), le sous-espace vectoriel :

$$F = \bigcap_{i=1}^{p} \operatorname{Ker}\varphi_i = \{x \in E \mid \forall\, i \in [\![1\,;p]\!],\ \varphi_i(x) = 0\}$$

est un sous-espace vectoriel de E de dimension $n-p$.

L'ensemble des p équations $\varphi_i(x) = 0$ $(1 \leqslant i \leqslant p)$ s'appelle un système d'équations de F.

[S2.21] Projections, projecteurs, symétries

- **Définitions géométriques**

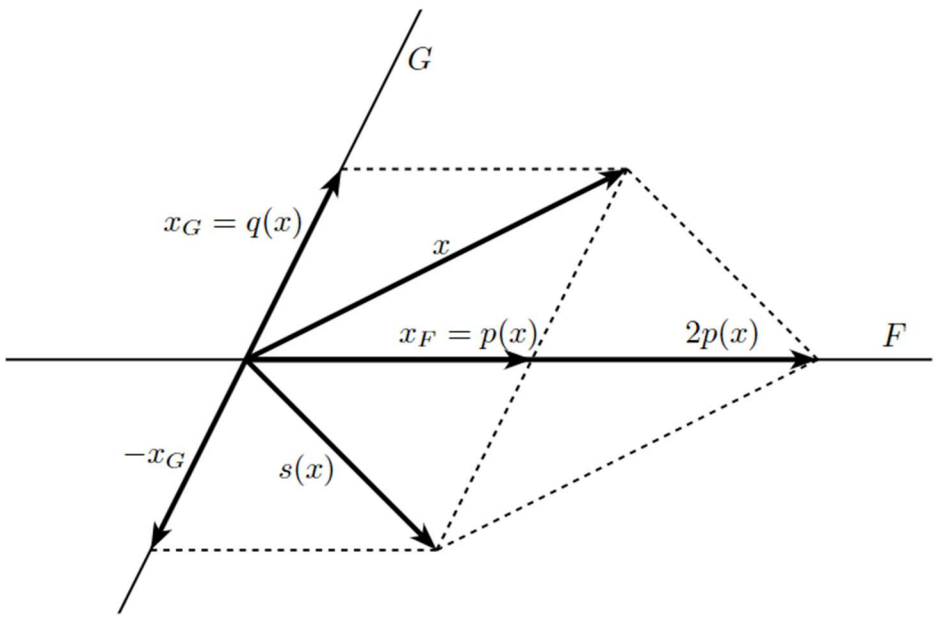
  Soit E un $\mathbb{K}$-espace vectoriel, et F, G deux sous-espaces vectoriels supplémentaires de E : $E = F \oplus G$.

  Tout vecteur $x \in E$ s'écrit donc de manière unique sous la forme : $x = x_F + x_G$ avec $(x_F, x_G) \in F \times G$.

  La projection sur F parallèlement à G est l'application $p : \begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x_F \end{cases}$.

  La symétrie par rapport à F et parallèlement à G est l'application

  $s : \begin{cases} E & \longrightarrow & E \\ x & \longmapsto & x_F - x_G \end{cases}$.



- **Propriétés**
  - $s = 2p - \mathrm{Id}_E$.
  - s et p sont deux endomorphismes de E.
  - $p^2 = p \circ p = p$.
  - $s^2 = s \circ s = \mathrm{Id}_E$ (s est donc bijectif, et $s^{-1} = s$).
  - $\mathrm{Ker}(p - \mathrm{Id}_E) = \mathrm{Im}\, p = F$ et $\mathrm{Ker}\, p = G$.
  - $\mathrm{Ker}(s - \mathrm{Id}_E) = F$ et $\mathrm{Ker}(s + \mathrm{Id}_E) = G$.
  - Si p est la projection sur F parallèlement à G, et q est la projection sur G parallèlement à F, $p + q = \mathrm{Id}_E$ (p et q sont dits associés).
- **Projecteurs**

  On appelle projecteur de E tout endomorphisme p de E tel que $p^2 = p$.

Si p est un projecteur, alors :

$$E = \operatorname{Im} p \oplus \operatorname{Ker} p \quad ; \quad \operatorname{Im} p = \operatorname{Ker}(p - \operatorname{Id}_E)$$

et p est la projection sur $\operatorname{Im} p$ parallèlement à $\operatorname{Ker} p$.

✓ Pour un projecteur p, il est vraiment important de retenir la propriété $\operatorname{Im} p = \operatorname{Ker}(p - \operatorname{Id}_E)$, qui exprime que les vecteurs de l'image de p sont exactement les vecteurs invariants par p.

✓ Il est également utile de retenir que la décomposition de tout vecteur $x \in E$ selon les sous-espaces vectoriels supplémentaires $\operatorname{Im} p$ et $\operatorname{Ker} p$ est :

$$x = \underbrace{p(x)}_{\in \operatorname{Im} p} + \underbrace{x - p(x)}_{\in \operatorname{Ker} p} .$$

- **Endomorphismes involutifs**

Si s est un endomorphisme involutif de E ($s^2 = \operatorname{Id}_E$), alors :

$$E = \operatorname{Ker}(s - \operatorname{Id}_E) \oplus \operatorname{Ker}(s + \operatorname{Id}_E) ,$$

et s est la symétrie par rapport à $\operatorname{Ker}(s - \operatorname{Id}_E)$ parallèlement à $\operatorname{Ker}(s + \operatorname{Id}_E)$.

✓ Il peut être utile de retenir la décomposition suivante :

$$\forall x \in E,\ x = \underbrace{\frac{x + s(x)}{2}}_{\in \operatorname{Ker}(s - \operatorname{Id}_E)} + \underbrace{\frac{x - s(x)}{2}}_{\in \operatorname{Ker}(s + \operatorname{Id}_E)} .$$

- **Projecteurs associés à une décomposition en somme directe**

Soit $(E_i)_{1 \leqslant i \leqslant n}$ une famille de sous-espaces vectoriels de E telle que $E = \bigoplus_{i=1}^{n} E_i$.

Pour tout $x \in E$, il existe donc une unique famille $(x_i)_{1 \leqslant i \leqslant n} \in \prod_{i=1}^{n} E_i$ telle que : $x = \sum_{i=1}^{n} x_i$.

Considérons, pour tout $i \in [\![1\,;n]\!]$, l'application $p_i : \begin{cases} E & \to & E \\ x & \mapsto & x_i \end{cases}$.

On a alors :

– $\forall i \in [\![1\,;n]\!]$, $p_i \circ p_i = p_i$.

– $\forall i \in [\![1\,;n]\!]$, p_i est la projection sur E_i parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq i}}^{n} E_j$.

– $\sum_{i=1}^{n} p_i = \operatorname{Id}_E$.

– $\forall (i,j) \in [\![1\,;n]\!]^2,\ i \neq j \implies p_i \circ p_j = 0_{\mathscr{L}(E)}$.

On dit que $(p_i)_{1 \leqslant i \leqslant n}$ est la famille de projecteurs canoniquement associée à la décomposition de E en somme directe $E = \bigoplus_{i=1}^{n} E_i$.

Thème 3 - Calcul matriciel

Dans toute la suite, $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$ (le plus souvent $\mathbb{R}$ ou $\mathbb{C}$), et p, q et n désignent des entiers naturels non nuls. On note :

- $\mathcal{M}_{p,q}(\mathbb{K})$ l'ensemble des matrices de type (p,q) à coefficients dans $\mathbb{K}$.
- $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n.

[S3.1] Opérations sur les matrices

- **Structure de $\mathbb{K}$-espace vectoriel**

  Soient $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B = (b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ deux matrices de même type (p,q), et soit $\lambda \in \mathbb{K}$ un scalaire.

  La somme des matrices A et B est définie par :

  $$A + B = (a_{ij} + b_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}},$$

  et la multiplication externe de A par λ est définie par :

  $$\lambda.A = (\lambda a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}.$$

  Muni de ces lois, $(\mathcal{M}_{p,q}(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel.

  L'élément neutre pour l'addition dans $\mathcal{M}_{p,q}(\mathbb{K})$ est la matrice nulle (dont tous les termes sont nuls), que l'on note $0_{p,q}$, ou 0 s'il n'y a pas d'ambiguïté.

- **Produit de matrices**

  Soient $A = (a_{ij})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{jk})_{\substack{1\leqslant j\leqslant p\\1\leqslant k\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ deux matrices.

  On définit le produit matriciel $AB = (c_{ik})_{\substack{1\leqslant i\leqslant n\\1\leqslant k\leqslant q}} \in \mathcal{M}_{n,q}(\mathbb{K})$ par :

  $$\forall\,(i,k) \in [\![1\,;n]\!] \times [\![1\,;q]\!],\ c_{ik} = \sum_{j=1}^{p} a_{ij}b_{jk}\,.$$

  Pour tout $k \in [\![1\,;q]\!]$, la k^{e} colonne de AB est égale à AC_k, où C_k désigne la k^{e} colonne de B. De même, pour tout $i \in [\![1\,;n]\!]$, la i^{e} ligne de AB est égale à L_iB, où L_i désigne la i^{e} ligne de A.

  > ✓ Le produit matriciel AB n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B.

- **Propriétés**
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ \forall\, C \in \mathcal{M}_{r,s}(\mathbb{K}),\ A(BC) = (AB)C.$
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B_1,\ B_2 \in \mathcal{M}_{q,r}(\mathbb{K}),\ A(B_1 + B_2) = AB_1 + AB_2.$
  - $\forall\, A_1,\ A_2 \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ (A_1 + A_2)B = A_1B + A_2B.$
  - $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{q,r}(\mathbb{K}),\ \forall\, \lambda \in \mathbb{K},\ (\lambda A)B = A(\lambda B) = \lambda(AB).$

On peut résumer ces propriétés en disant que la multiplication des matrices est une opération associative et bilinéaire.

- **Base canonique de $\mathcal{M}_{p,q}(\mathbb{K})$**
  - Pour tout $(k,\ell) \in [\![1\,;p]\!] \times [\![1\,;q]\!]$, on note $E_{k\ell}$ la matrice de type (p,q) dont tous les termes sont nuls, sauf celui d'indice (k,ℓ) qui vaut 1. Autrement dit,
  $$E_{k\ell} = (\delta_{ki}\delta_{\ell j})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})\,.$$
  - La famille des matrices $(E_{k\ell})_{(k,\ell)\in[\![1;p]\!]\times[\![1;q]\!]}$ forme une base de $\mathcal{M}_{p,q}(\mathbb{K})$, appelée base canonique.
    On en déduit que : $\dim(\mathcal{M}_{p,q}(\mathbb{K})) = pq$.
  - Pour toute matrice $A = (a_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$, $A = \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} a_{ij}E_{ij}$, c'est-à-dire que les coordonnées de A dans la base canonique sont les a_{ij}.
- Soient $(E_{ij})_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}$, $(E'_{jk})_{\substack{1\leqslant j\leqslant q\\1\leqslant k\leqslant r}}$ et $(E''_{ik})_{\substack{1\leqslant i\leqslant p\\1\leqslant k\leqslant r}}$ les bases canoniques respectives de $\mathcal{M}_{p,q}(\mathbb{K})$, de $\mathcal{M}_{q,r}(\mathbb{K})$ et de $\mathcal{M}_{p,r}(\mathbb{K})$. Alors :
$$\forall\,(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!],\ \forall\,(k,\ell) \in [\![1\,;q]\!] \times [\![1\,;r]\!],\ E_{ij}E'_{k\ell} = \delta_{jk}E''_{i\ell}\,.$$

[S3.2] Opérations par blocs

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. On appelle bloc de A toute matrice $(a_{ij})_{\substack{i\in I\\j\in J}}$, où I et J sont respectivement des parties de $[\![1\,;p]\!]$ et de $[\![1\,;q]\!]$ formées d'entiers *consécutifs* (lorsque les indices ne sont pas consécutifs on parle de matrice *extraite*).
- **Combinaison linéaire par blocs**
  Soient A, $B \in \mathcal{M}_{p,q}(\mathbb{K})$ décomposées en blocs avec le même découpage :
$$A = \underbrace{\left[\begin{array}{ccc} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{array}\right]}_{q}\begin{array}{l}\updownarrow p_1\\ \\ \updownarrow p_\ell\end{array}\Big\uparrow p \quad \text{et} \quad B = \underbrace{\left[\begin{array}{ccc} B_{11} & \dots & B_{1m} \\ \vdots & & \vdots \\ B_{\ell 1} & \dots & B_{\ell m} \end{array}\right]}_{q}\begin{array}{l}\updownarrow p_1\\ \\ \updownarrow p_\ell\end{array}\Big\uparrow p\,.$$
  Alors, pour tout $\lambda \in \mathbb{K}$, la matrice $\lambda A + B$ s'écrit, par blocs :
$$\lambda A + B = \underbrace{\left[\begin{array}{ccc} \lambda A_{11} + B_{11} & \dots & \lambda A_{1m} + B_{1m} \\ \vdots & & \vdots \\ \lambda A_{\ell 1} + B_{\ell 1} & \dots & \lambda A_{\ell m} + B_{\ell m} \end{array}\right]}_{q}\begin{array}{l}\updownarrow p_1\\ \\ \updownarrow p_\ell\end{array}\Big\uparrow p\,.$$

- **Produit par blocs**

Soient $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$, décomposées en blocs comme suit :

$$A = \underbrace{\overbrace{\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{bmatrix}}^{\begin{matrix} q_1 & & q_m \end{matrix}}}_{q} \begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix} \Big\} p \quad \text{et} \quad B = \underbrace{\overbrace{\begin{bmatrix} B_{11} & \dots & B_{1n} \\ \vdots & & \vdots \\ B_{m1} & \dots & B_{mn} \end{bmatrix}}^{\begin{matrix} r_1 & & r_n \end{matrix}}}_{r} \begin{matrix} \updownarrow q_1 \\ \\ \updownarrow q_m \end{matrix} \Big\} q,$$

et soit $C = AB \in \mathcal{M}_{p,r}(\mathbb{K})$, décomposée en blocs :

$$C = \underbrace{\overbrace{\begin{bmatrix} C_{11} & \dots & C_{1n} \\ \vdots & & \vdots \\ C_{\ell 1} & \dots & C_{\ell n} \end{bmatrix}}^{\begin{matrix} r_1 & & r_n \end{matrix}}}_{r} \begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix} \Big\} p\,.$$

Alors : $\forall\,(i,k) \in [\![1\,;\ell]\!] \times [\![1\,;n]\!]$, $C_{ik} = \sum_{j=1}^{m} A_{ij}B_{jk}$.

[S3.3] Transposition

Soit $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$.

On appelle transposée de A la matrice ${}^{\mathrm{t}}A = (a'_{ji})_{\substack{1 \leqslant j \leqslant q \\ 1 \leqslant i \leqslant p}} \in \mathcal{M}_{q,p}(\mathbb{K})$ définie par :

$$\forall\,(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!],\ a'_{ji} = a_{ij}\,.$$

La transposée de A peut aussi être notée A^T.

Propriétés

- $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, ${}^{\mathrm{t}}({}^{\mathrm{t}}A) = A$.
- $\forall\, A,\, B \in \mathcal{M}_{p,q}(\mathbb{K})$, ${}^{\mathrm{t}}(A+B) = {}^{\mathrm{t}}A + {}^{\mathrm{t}}B$.
- $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, $\forall\, \lambda \in \mathbb{K}$, ${}^{\mathrm{t}}(\lambda A) = \lambda\, {}^{\mathrm{t}}A$.
- $\forall\, A \in \mathcal{M}_{p,q}(\mathbb{K})$, $\forall\, B \in \mathcal{M}_{q,r}(\mathbb{K})$, ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\, {}^{\mathrm{t}}A$.
- L'application $\left\{ \begin{array}{ccc} \mathcal{M}_{p,q}(\mathbb{K}) & \longrightarrow & \mathcal{M}_{q,p}(\mathbb{K}) \\ A & \longmapsto & {}^{\mathrm{t}}A \end{array} \right.$ est un isomorphisme d'espaces vectoriels.

Transposition par blocs

Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$.

$$\text{Si } A = \underbrace{\overbrace{\begin{bmatrix} A_{11} & \dots & A_{1m} \\ \vdots & & \vdots \\ A_{\ell 1} & \dots & A_{\ell m} \end{bmatrix}}^{\begin{matrix} q_1 & & q_m \end{matrix}}}_{q} \begin{matrix} \updownarrow p_1 \\ \\ \updownarrow p_\ell \end{matrix} \Big\} p \quad \text{alors} \quad {}^{\mathrm{t}}A = \underbrace{\overbrace{\begin{bmatrix} {}^{\mathrm{t}}A_{11} & \dots & {}^{\mathrm{t}}A_{\ell 1} \\ \vdots & & \vdots \\ {}^{\mathrm{t}}A_{1m} & \dots & {}^{\mathrm{t}}A_{\ell m} \end{bmatrix}}^{\begin{matrix} p_1 & & p_\ell \end{matrix}}}_{p} \begin{matrix} \updownarrow q_1 \\ \\ \updownarrow q_m \end{matrix} \Big\} q\,.$$

[S3.4] L'algèbre des matrices carrées

- On sait que $(\mathcal{M}_n(\mathbb{K}), +, .)$ est un $\mathbb{K}$-espace vectoriel de dimension n^2. De plus, dans $\mathcal{M}_n(\mathbb{K})$, la multiplication matricielle $\times$ est une loi de composition interne, et $(\mathcal{M}_n(\mathbb{K}), +, ., \times)$ est une $\mathbb{K}$-algèbre.

  L'élément neutre pour la loi $\times$ est la matrice identité I_n dont les termes diagonaux valent 1, et dont les termes non diagonaux sont nuls.

  L'algèbre $\mathcal{M}_n(\mathbb{K})$ n'est pas commutative et n'est pas intègre dès que $n \geqslant 2$.

- **Matrices inversibles**

  On note $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles d'ordre n. Il s'agit d'un groupe pour la multiplication matricielle, appelé groupe linéaire d'ordre n. C'est le groupe des éléments inversibles de l'anneau $(\mathcal{M}_n(\mathbb{K}), +, \times)$.

  Si A est inversible, il en est de même de ${}^tA$, et $({}^tA)^{-1} = {}^t(A^{-1})$.

- Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors les propriétés suivantes sont équivalentes :

  (i) A est inversible,

  (ii) A est inversible à droite $(\exists\, B \in \mathcal{M}_n(\mathbb{K}) \mid AB = \mathrm{I}_n)$,

  (iii) A est inversible à gauche $(\exists\, B \in \mathcal{M}_n(\mathbb{K}) \mid BA = \mathrm{I}_n)$,

  (iv) A est régulière à droite $(\forall\, M, N \in \mathcal{M}_n(\mathbb{K}),\ MA = NA \implies M = N)$,

  (v) A est régulière à gauche $(\forall\, M, N \in \mathcal{M}_n(\mathbb{K}),\ AM = AN \implies M = N)$.

[S3.5] Matrices carrées remarquables

- **Matrices triangulaires**

  . Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. A est dite triangulaire supérieure (respectivement inférieure) si pour tout $(i,j) \in [\![1\,;n]\!]^2$:

  $$i > j \implies a_{ij} = 0 \quad (\text{respectivement } i < j \implies a_{ij} = 0)\,.$$

  On note respectivement $\mathcal{T}_n^+(\mathbb{K})$ et $\mathcal{T}_n^-(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures et l'ensemble des matrices triangulaires inférieures d'ordre n.

  . **Propriétés**

  – Soient $A = (a_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$ deux matrices triangulaires supérieures. Alors $C = AB$ est triangulaire supérieure et ses coefficients diagonaux sont les scalaires $c_{ii} = a_{ii}b_{ii}$.

  – Soit $A = (a_{ij}) \in \mathcal{T}_n^+(\mathbb{K})$. Alors :

  $$A \text{ inversible} \iff \forall\, i \in [\![1\,;n]\!],\ a_{ii} \neq 0\,.$$

  De plus, dans ce cas, A^{-1} est triangulaire supérieure et ses termes diagonaux sont les scalaires $\dfrac{1}{a_{ii}}$.

  – $\mathcal{T}_n^+(\mathbb{K})$ est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $\dfrac{n(n+1)}{2}$.

– L'application $\left\{\begin{array}{ccc}\mathcal{T}_n^+(\mathbb{K}) & \longrightarrow & \mathcal{T}_n^-(\mathbb{K}) \\ A & \longmapsto & {}^tA\end{array}\right.$ est un anti-isomorphisme d'algèbres.

Il en résulte que les propriétés précédentes (produit, inverse et dimension) sont encore vraies pour les matrices triangulaires inférieures.

- **Matrices diagonales**

. Une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est dite diagonale si :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ i \neq j \implies a_{ij} = 0\,.$$

On note $\mathcal{D}_n(\mathbb{K}) = \mathcal{T}_n^+(\mathbb{K}) \cap \mathcal{T}_n^-(\mathbb{K})$ l'ensemble des matrices diagonales d'ordre n. $\mathcal{D}_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$ et $\dim \mathcal{D}_n(\mathbb{K}) = n$.
Pour tout $(\lambda_1,\ldots,\lambda_n) \in \mathbb{K}^n$, on note $\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ la matrice diagonale $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ telle que : $\forall\, i \in [\![1\,;n]\!],\ a_{ii} = \lambda_i\,.$

. **Propriétés**

– Soient $A = \mathrm{diag}(a_{11},\ldots,a_{nn}) \in \mathcal{D}_n(\mathbb{K})$ et $B = \mathrm{diag}(b_{11},\ldots,b_{nn}) \in \mathcal{D}_n(\mathbb{K})$ deux matrices diagonales. Alors $AB = \mathrm{diag}(a_{11}b_{11},\ldots,a_{nn}b_{nn})$.

– Soit $A = \mathrm{diag}(a_{11},\ldots,a_{nn}) \in \mathcal{D}_n(\mathbb{K})$. Alors :

$$A \text{ inversible} \iff \forall\, i \in [\![1\,;n]\!],\ a_{ii} \neq 0\,.$$

De plus, dans ce cas, $A^{-1} = \mathrm{diag}\left(\dfrac{1}{a_{11}},\ldots,\dfrac{1}{a_{nn}}\right)$.

– Soit $D \in \mathcal{D}_n(\mathbb{K})$ une matrice diagonale dont les termes diagonaux sont *deux à deux distincts*. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

$$AD = DA \iff A \text{ diagonale}\,.$$

✓ Autrement dit, si une matrice commute avec une matrice diagonale *à éléments diagonaux distincts*, elle est diagonale. Ce résultat, très utile, ne figure pas au programme officiel et est démontré en **[SF3.1]**.

- **Matrices scalaires**

. Une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est dite scalaire s'il existe $\lambda \in \mathbb{K}$ tel que $A = \mathrm{diag}(\lambda,\ldots,\lambda) = \lambda \mathrm{I}_n$.

. Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

$$A \text{ scalaire} \iff \forall\, M \in \mathcal{M}_n(\mathbb{K}),\ AM = MA\,.$$

✓ Autrement dit, les matrices carrées qui commutent avec toutes les autres sont les matrices scalaires. Ce résultat classique ne figure pas au programme officiel et est démontré en **[SF3.2]**.

- **Matrices symétriques, matrices antisymétriques**

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée.
A est dite symétrique si ${}^tA = A$, et antisymétrique si ${}^tA = -A$.

✓ Les éléments diagonaux d'une matrice antisymétrique sont nécessairement nuls (c'est une conséquence de la définition).

On note respectivement $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices symétriques et l'ensemble des matrices antisymétriques d'ordre n.
Ce sont des sous-espaces vectoriels supplémentaires de $\mathcal{M}_n(\mathbb{K})$:

$$\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K}) .$$

Plus précisément, la décomposition unique d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est :

$$A = \underbrace{\frac{1}{2}(A + {}^{\mathrm{t}}A)}_{\in \mathcal{S}_n(\mathbb{K})} + \underbrace{\frac{1}{2}(A - {}^{\mathrm{t}}A)}_{\in \mathcal{A}_n(\mathbb{K})} .$$

Enfin, on a : $\dim\big(\mathcal{S}_n(\mathbb{K})\big) = \dfrac{n(n+1)}{2}$ et $\dim\big(\mathcal{A}_n(\mathbb{K})\big) = \dfrac{n(n-1)}{2}$.

[S3.6] Matrice d'une famille de vecteurs

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, rapporté à une base $\mathscr{B} = (e_1, \dots, e_n)$ et soient $x_1, \dots, x_p$ des vecteurs de E.

Pour tout $j \in [\![1\,;p]\!]$, le vecteur x_j s'écrit dans la base $\mathscr{B}$ sous la forme $x_j = \sum_{i=1}^{n} a_{ij} e_i$ avec $a_{ij} \in \mathbb{K}$.
La matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} \in \mathcal{M}_{n,p}(\mathbb{K})$ s'appelle la matrice de la famille $(x_1, \dots, x_p)$ dans la base $\mathscr{B}$.
Il s'agit donc de la matrice obtenue en écrivant dans chaque colonne les coordonnées dans $\mathscr{B}$ des vecteurs de la famille.

[S3.7] Matrice d'une application linéaire

- Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension q, rapporté à une base $\mathscr{B}_E = (e_1, \dots, e_q)$;
  - F un $\mathbb{K}$-espace vectoriel de dimension p, rapporté à une base $\mathscr{B}_F = (e'_1, \dots, e'_p)$;
  - $u \in \mathscr{L}(E,F)$.

  On appelle matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$ la matrice dans $\mathscr{B}_F$ du système de vecteurs $\big(u(e_1), \dots, u(e_q)\big)$. On la note généralement $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$.
  Ainsi : $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = (a_{ij}) \in \mathcal{M}_{p,q}(\mathbb{K})$, où a_{ij} désigne la i^{e} coordonnée de $u(e_j)$ dans la base $\mathscr{B}_F$ soit :

$$\forall\, j \in [\![1\,;q]\!] \ , \ u(e_j) = \sum_{i=1}^{p} a_{ij} e'_i .$$

  L'écriture de la matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$ peut être visualisée ainsi :

$$M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \dots & u(e_q) \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pq} \end{pmatrix} \end{array} \begin{array}{l} \\ \to e'_1 \\ \to e'_2 \\ \vdots \\ \to e'_p \end{array}$$

✓ Il est important de remarquer que si u est une application linéaire d'un espace de dimension q vers un espace de dimension p alors $M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est une matrice de type (p,q).

- L'application $\begin{cases} \mathscr{L}(E,F) & \longrightarrow & \mathcal{M}_{p,q}(\mathbb{K}) \\ u & \longmapsto & M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) \end{cases}$ est un isomorphisme d'espaces vectoriels, c'est-à-dire qu'elle est bijective et que :

$$\forall\,(u,v) \in \mathscr{L}(E,F)^2,\ \forall\,\lambda \in \mathbb{K},\ M_{\mathscr{B}_E}^{\mathscr{B}_F}(\lambda u + v) = \lambda . M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) + M_{\mathscr{B}_E}^{\mathscr{B}_F}(v)\,.$$

En particulier, pour toute matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$, il existe une unique application linéaire $a \in \mathscr{L}(\mathbb{K}^q,\mathbb{K}^p)$ telle que la matrice de a dans les bases canoniques de $\mathbb{K}^q$ et $\mathbb{K}^p$ soit égale à A. L'application a ainsi définie est appelée l'application linéaire canoniquement associée à A.

- **Matrice d'un endomorphisme**

Dans le cas où u est un endomorphisme d'un espace vectoriel E de dimension n muni d'une base $\mathscr{B}$, on appelle matrice de u dans la base $\mathscr{B}$ la matrice $M_{\mathscr{B}}^{\mathscr{B}}(u) \in \mathcal{M}_n(\mathbb{K})$, et on la note simplement $M_{\mathscr{B}}(u)$.
Pour toute base $\mathscr{B}$ de E, on a $M_{\mathscr{B}}(\mathrm{Id}_E) = \mathrm{I}_n$. Plus généralement, si u est une homothétie de E de rapport λ, $M_{\mathscr{B}}(u) = \lambda \mathrm{I}_n$: c'est une matrice scalaire.

- **Matrice d'une composée d'applications linéaires**
  - Soient E, F et G des $\mathbb{K}$-espaces vectoriels de bases respectives $\mathscr{B}_E$, $\mathscr{B}_F$ et $\mathscr{B}_G$. Soient $u \in \mathscr{L}(E,F)$ et $v \in \mathscr{L}(F,G)$ deux applications linéaires. Alors :
$$M_{\mathscr{B}_E}^{\mathscr{B}_G}(v \circ u) = M_{\mathscr{B}_F}^{\mathscr{B}_G}(v) \times M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)\,.$$
  - Soit E un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathscr{B}$, et soit $u \in \mathscr{L}(E)$. On note $A = M_{\mathscr{B}}(u)$. Alors :
$$u \in \mathrm{GL}(E) \iff A \in \mathrm{GL}_n(\mathbb{K})\,.$$
    De plus, dans ce cas, $A^{-1} = M_{\mathscr{B}}(u^{-1})$.

- **Expression analytique d'une application linéaire**

Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension q, rapporté à une base $\mathscr{B}_E = (e_1,\ldots,e_q)$;
  - F un $\mathbb{K}$-espace vectoriel de dimension p, rapporté à une base $\mathscr{B}_F = (e'_1,\ldots,e'_p)$;
  - $u \in \mathscr{L}(E,F)$;
  - $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u) = (a_{ij}) \in \mathcal{M}_{p,q}(\mathbb{K})$ la matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$.

Soit $x = \sum_{j=1}^{q} x_j e_j \in E$, et soit $y = u(x) = \sum_{i=1}^{p} y_i e'_i \in F$ son image par u. Alors :

$$\forall\, i \in [\![1\,;p]\!],\ y_i = \sum_{j=1}^{q} a_{ij} x_j \quad \text{(expression analytique de } u \text{ dans } \mathscr{B}_E \text{ et } \mathscr{B}_F\text{)}.$$

Cela s'écrit matriciellement : $\boxed{Y = AX}$, où :
- $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \ldots & x_q \end{pmatrix}$ est la matrice colonne des coordonnées de x dans $\mathscr{B}_E$;
- $Y = {}^{\mathrm{t}}\begin{pmatrix} y_1 & \ldots & y_p \end{pmatrix}$ est la matrice colonne des coordonnées de y dans $\mathscr{B}_F$.

[S3.8] Matrices par blocs et sous-espaces stables

- Soit E un $\mathbb{K}$-espace vectoriel, et u un endomorphisme de E.
  Un sous-espace vectoriel F de E est dit stable par u si $u(F) \subset F$, c'est-à-dire si :
  $$\forall\, x \in F,\ u(x) \in F\,.$$
  On peut alors définir l'endomorphisme u_F induit par u sur F par :
  $$\forall\, x \in F,\ u_F(x) = u(x)\,.$$

  ✓ Il ne faut pas confondre l'endomorphisme induit u_F, qui n'est défini que si F est stable par u et qui est une application linéaire de F dans F, avec la restriction $u|_F$, qui est une application linéaire de F dans E.

- Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$ et F un sous-espace de E, de dimension p. Soit $\mathscr{B}_F$ une base de F, que l'on complète en une base $\mathscr{B}_E$ de E.
  Soit u un endomorphisme de E et $A = M_{\mathscr{B}_E}(u)$ sa matrice dans $\mathscr{B}_E$.
  Alors F est stable par u si et seulement si A est de la forme :
  $$\begin{bmatrix} A_1 & B \\ 0_{n-p,p} & A_2 \end{bmatrix} \quad \text{avec } A_1 \in \mathcal{M}_p(\mathbb{K}).$$
  Une telle matrice est dite triangulaire supérieure par blocs.
  Dans ce cas, $A_1 = M_{\mathscr{B}_F}(u_F)$ où u_F est l'endomorphisme induit par u sur F.

- **Généralisation**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$, $E_1, \ldots, E_p$ p sous-espaces vectoriels supplémentaires de E et $\mathscr{B}_E = \bigcup_{i=1}^{p} \mathscr{B}_i$ une base adaptée à la décomposition en somme directe $E = \bigoplus_{i=1}^{p} E_i$.
  Soit u un endomorphisme de E. Alors les sous-espaces E_i sont stables par u si et seulement si la matrice A de u dans $\mathscr{B}_E$ est diagonale par blocs, c'est-à-dire de la forme :
  $$A = \begin{bmatrix} A_1 & 0 & & 0 \\ 0 & A_2 & \ddots & \\ & \ddots & \ddots & 0 \\ 0 & & 0 & A_p \end{bmatrix} \quad \text{avec } A_i \in \mathcal{M}_{n_i}(\mathbb{K}) \text{ où } n_i = \dim E_i\,.$$
  Pour tout $i \in [\![1\,;p]\!]$, A_i est alors la matrice dans $\mathscr{B}_i$ de l'endomorphisme u_i induit par u sur E_i.

[S3.9] Rang d'une matrice

Le rang d'une matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$ est, par définition, le rang de ses vecteurs colonnes (éléments de $\mathcal{M}_{p,1}(\mathbb{K})$, que l'on peut assimiler à des éléments de $\mathbb{K}^p$). On le note $\operatorname{rg} A$. D'autres interprétations sont possibles :

– si $A = M_{\mathscr{B}_F}(x_1,\dots,x_q)$ est la matrice d'une famille de vecteurs $(x_1,\dots,x_q)$ de F relativement à une base $\mathscr{B}_F$, alors :

$$\operatorname{rg} A = \operatorname{rg}(x_1,\dots,x_q) = \dim\big(\operatorname{Vect}(x_1,\dots,x_q)\big)\,;$$

– si $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est la matrice d'une application linéaire $u \in \mathscr{L}(E,F)$ relativement à des bases $\mathscr{B}_E$ et $\mathscr{B}_F$, alors $\operatorname{rg} A = \operatorname{rg} u$.

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors on a les résultats suivants.
  – $\operatorname{rg} A \leqslant \min(p,q)$.
  – $\operatorname{rg} A = \operatorname{rg}({}^{\mathrm{t}}A)$.
  – Si $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ est la matrice d'une application linéaire $u \in \mathscr{L}(E,F)$ dans des bases $\mathscr{B}_E$ et $\mathscr{B}_F$, alors :

$$\operatorname{rg} A = p \iff u \text{ surjective} \quad ; \quad \operatorname{rg} A = q \iff u \text{ injective}.$$

- En particulier, pour $A \in \mathcal{M}_n(\mathbb{K})$, on a :

$$A \text{ inversible} \iff \operatorname{rg} A = n\,.$$

- **Rang d'un produit**
  Soient $A \in \mathcal{M}_{p,q}(\mathbb{K})$ et $B \in \mathcal{M}_{q,r}(\mathbb{K})$. On a les résultats suivants.
  – $\operatorname{rg}(AB) \leqslant \min(\operatorname{rg} A, \operatorname{rg} B)$.
  – Si $p = q$ et $A \in \mathrm{GL}_p(\mathbb{K})$, alors $\operatorname{rg}(AB) = \operatorname{rg}(B)$.
  – Si $q = r$ et $B \in \mathrm{GL}_q(\mathbb{K})$, alors $\operatorname{rg}(AB) = \operatorname{rg}(A)$.

- **Caractérisation des matrices de rang r**
  Le rang d'une matrice non nulle $A \in \mathcal{M}_{p,q}(\mathbb{K})$ est le plus grand entier $r \geqslant 1$ tel que l'on puisse extraire de A une matrice carrée inversible d'ordre r.

[S3.10] Formules de changement de base

- **Matrices de passage**
  . Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et $\mathscr{B} = (e_1,\dots,e_n)$ et $\mathscr{B}' = (e'_1,\dots,e'_n)$ deux bases de E.
  On appelle matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ la matrice du système $(e'_1,\dots,e'_n)$ dans la base $\mathscr{B}$. On la note généralement $P_{\mathscr{B}}^{\mathscr{B}'}$.

  ✓ Ainsi, pour former la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, on écrit dans les colonnes les coordonnées dans $\mathscr{B}$ des vecteurs de $\mathscr{B}'$.

  . **Propriétés**
  – Si $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ sont trois bases de E, alors $P_{\mathscr{B}}^{\mathscr{B}''} = P_{\mathscr{B}}^{\mathscr{B}'} \times P_{\mathscr{B}'}^{\mathscr{B}''}$.
  – $P_{\mathscr{B}}^{\mathscr{B}'}$ est inversible et $\left(P_{\mathscr{B}}^{\mathscr{B}'}\right)^{-1} = P_{\mathscr{B}'}^{\mathscr{B}}$.

- **Formule de changement de base pour un vecteur**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, et P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.
  Soit x un vecteur de E, X la matrice colonne de ses coordonnées dans $\mathscr{B}$ et X' celle de ses coordonnées dans $\mathscr{B}'$.
  On a alors la relation : $\boxed{X = PX'}$.
- **Formule de changement de bases pour une application linéaire**
  Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension q, muni de deux bases $\mathscr{B}_E$ et $\mathscr{B}'_E$;
  - F un $\mathbb{K}$-espace vectoriel de dimension p, muni de deux bases $\mathscr{B}_F$ et $\mathscr{B}'_F$;
  - $P \in \mathrm{GL}_q(\mathbb{K})$ la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$;
  - $Q \in \mathrm{GL}_p(\mathbb{K})$ la matrice de passage de $\mathscr{B}_F$ à $\mathscr{B}'_F$;
  - $u \in \mathscr{L}(E,F)$, $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$ la matrice de u dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$, $A' = M_{\mathscr{B}'_E}^{\mathscr{B}'_F}(u)$ celle dans les bases $\mathscr{B}'_E$ et $\mathscr{B}'_F$ (A, $A' \in \mathcal{M}_{p,q}(\mathbb{K})$).

  On a alors la relation : $\boxed{A' = Q^{-1}AP}$.
- **Formule de changement de base pour un endomorphisme**
  Soient :
  - E un $\mathbb{K}$-espace vectoriel de dimension n, muni de deux bases $\mathscr{B}_E$ et $\mathscr{B}'_E$;
  - $P \in \mathrm{GL}_n(\mathbb{K})$ la matrice de passage de $\mathscr{B}_E$ à $\mathscr{B}'_E$;
  - $u \in \mathscr{L}(E)$, $A = M_{\mathscr{B}_E}(u)$ sa matrice dans la base $\mathscr{B}$ et $A' = M_{\mathscr{B}'_E}(u)$ celle dans la base $\mathscr{B}'$ (A, $A' \in \mathcal{M}_n(\mathbb{K})$).

  On a alors la relation : $\boxed{A' = P^{-1}AP}$.

[S3.11] Matrices équivalentes

- On dit qu'une matrice $A' \in \mathcal{M}_{p,q}(\mathbb{K})$ est équivalente à la matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$ s'il existe $Q \in \mathrm{GL}_p(\mathbb{K})$ et $P \in \mathrm{GL}_q(\mathbb{K})$ telles que $A' = Q^{-1}AP$.
  - On vérifie facilement qu'il s'agit bien d'une relation d'équivalence dans $\mathcal{M}_{p,q}(\mathbb{K})$.
  - A et A' sont équivalentes si et seulement si ce sont les matrices, dans des couples de bases différentes, d'une même application linéaire u d'un espace vectoriel de dimension q dans un espace vectoriel de dimension p.
- **Caractérisation des matrices de rang r**
  Une matrice $A \in \mathcal{M}_{p,q}(\mathbb{K})$ est de rang r si et seulement si elle est équivalente à la matrice :
  $$J_{p,q,r} = \begin{bmatrix} I_r & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{bmatrix}.$$
- Deux matrices A et $A' \in \mathcal{M}_{p,q}(\mathbb{K})$ sont équivalentes si et seulement si elles ont même rang.

[S3.12] Matrices semblables, trace

- **Matrices semblables**
  Soient A et A' deux matrices carrées de $\mathcal{M}_n(\mathbb{K})$. On dit que A et A' sont semblables s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A' = P^{-1}AP$.

  ✓ Dire que A et A' sont semblables signifie donc que ce sont les matrices dans deux bases différentes d'un même endomorphisme u d'un espace vectoriel de dimension n. Cette interprétation est importante à retenir.

- On appelle trace d'une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire :

$$\mathrm{tr}\, A \;=\; \sum_{i=1}^{n} a_{ii} \quad \text{(somme des éléments diagonaux)}\,.$$

  – L'application $\mathrm{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est une forme linéaire.
  – $\forall\, A \in \mathcal{M}_{n,p}(\mathbb{K}),\ \forall\, B \in \mathcal{M}_{p,n}(\mathbb{K}),\ \mathrm{tr}(AB) = \mathrm{tr}(BA)$.
  – Deux matrices semblables ont même trace.

- **Trace d'un endomorphisme**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$.
  Pour toute base $\mathscr{B}$ de E, le scalaire $\mathrm{tr}(M_{\mathscr{B}}(u))$ ne dépend pas de la base $\mathscr{B}$ choisie. Ce scalaire s'appelle la trace de l'endomorphisme u, notée $\mathrm{tr}\, u$.

  – L'application $\mathrm{tr} : \mathscr{L}(E) \to \mathbb{K}$ est une forme linéaire.
  – $\forall\, u,\, v \in \mathscr{L}(E),\ \mathrm{tr}(v \circ u) = \mathrm{tr}(u \circ v)$.

- **Trace d'un projecteur**
  Si p est un projecteur de E, alors $\mathrm{tr}\, p = \mathrm{rg}\, p$.

[S3.13] Opérations élémentaires sur les lignes d'une matrice

- Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$, de lignes notées L_i $(1 \leqslant i \leqslant p)$ et de colonnes notées C_j $(1 \leqslant j \leqslant q)$.
  On appelle opération élémentaire sur les lignes de A l'une des opérations suivantes :

  – échange des lignes L_i et L_j, notée $L_i \leftrightarrow L_j$;
  – ajout à la ligne L_i de la ligne L_j multipliée par un scalaire $\lambda \in \mathbb{K}$, où $i \neq j$, notée $L_i \leftarrow L_i + \lambda L_j$;
  – multiplication de la ligne L_i par un scalaire $\lambda \in \mathbb{K}^*$, notée $L_i \leftarrow \lambda L_i$.

  On définit de manière analogue les opérations élémentaires sur les colonnes.

- **Matrices élémentaires**
  Soit $(E_{ij})_{1 \leqslant i,j \leqslant p}$ la base canonique de $\mathcal{M}_p(\mathbb{K})$, et soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$.

  . Si A' est la matrice obtenue à partir de A par l'opération :

$$L_i \leftarrow L_i + \lambda L_j \ (i \neq j)\,,$$

  alors $A' = T_{i,j}(\lambda)A$, où $T_{i,j}(\lambda) = I_p + \lambda E_{ij}$ (matrice de transvection).

- Si A' est la matrice obtenue à partir de A par l'opération :

$$L_i \leftarrow \lambda L_i \ (\lambda \in \mathbb{K}^*)\,,$$

alors $A' = D_i(\lambda)A$, où $D_i(\lambda) = I_p + (\lambda - 1)E_{ii}$ (matrice de dilatation).

- Si A' est la matrice obtenue à partie de A par l'opération

$$L_i \leftrightarrow L_j\,,$$

alors $A' = P_{i,j}A$, où $P_{i,j} = I_p - E_{ii} - E_{jj} + E_{ij} + E_{ji}$ (matrice de transposition).

- Les matrices $T_{i,j}(\lambda)$ avec $i \neq j$, $D_i(\lambda)$ avec $\lambda \neq 0$ et $P_{i,j}$ sont appelées matrices élémentaires. Ce sont des matrices carrées inversibles.

  ✓ On obtient des résultats similaires pour les opérations élémentaires sur les colonnes : il faut alors multiplier A *à droite* par une matrice élémentaire $Q \in \mathrm{GL}_q(\mathbb{K})$.

  ✓ Si A' est obtenue à partir de A par une suite d'opérations élémentaires sur les lignes ou les colonnes, A et A' sont équivalentes.

- **Matrices échelonnées**

  - Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$. Pour tout $i \in [\![1\,;p]\!]$, on note :

$$m_i = \begin{cases} \min\{j \in [\![1\,;q]\!] \mid a_{ij} \neq 0\} & \text{si } L_i \neq 0, \\ q + i & \text{si } L_i = 0. \end{cases}$$

    Si la ligne L_i est non nulle, m_i désigne donc le numéro de colonne du premier terme non nul de L_i en partant de la gauche. Le terme a_{ij} correspondant s'appelle un pivot.

  - On dit que A est échelonnée par lignes si la suite finie $(m_i)_{1 \leqslant i \leqslant p}$ est strictement croissante.

    Ainsi, dans une matrice échelonnée par lignes, la position des pivots croît strictement avec i jusqu'aux lignes nulles éventuelles, et si une ligne est nulle, alors toutes les lignes suivantes sont nulles.

  - Une matrice A échelonnée par lignes est dite échelonnée réduite par lignes si A est nulle ou si tous les pivots de A valent 1 et sont les seuls termes non nuls de leur colonne.

  - On définit de manière analogue les notions de matrice échelonnée par colonnes et de matrice échelonnée réduite par colonnes.

- **Théorème de Gauss-Jordan**

  Soit $A \in \mathcal{M}_{p,q}(\mathbb{K})$.

  – Il existe une unique matrice $R \in \mathcal{M}_{p,q}(\mathbb{K})$ échelonnée réduite par lignes telle que $A \underset{L}{\sim} R$. De plus, il existe $U \in \mathrm{GL}_p(\mathbb{K})$, produit de matrices élémentaires, telle que $UA = R$.

  – Il existe une unique matrice $R' \in \mathcal{M}_{p,q}(\mathbb{K})$ échelonnée réduite par colonnes telle que $A \underset{C}{\sim} R'$. De plus, il existe $V \in \mathrm{GL}_q(\mathbb{K})$, produit de matrices élémentaires, telle que $AV = R'$.

✓ La méthode du pivot permet ainsi de retrouver le résultat mentionné en **[S3.11]** : toute matrice de $\mathcal{M}_{p,q}(\mathbb{K})$ de rang r est équivalente à $J_{p,q,r}$.

- **Utilisation des opérations élémentaires**

  . **Calcul du rang d'une matrice**

  Si deux matrices A, $A' \in \mathcal{M}_{p,q}(\mathbb{K})$ sont équivalentes, elles ont même rang.

  Pour déterminer le rang de A, il suffit donc de la rendre équivalente par une suite d'opérations élémentaires sur les lignes ou les colonnes à une matrice triangulaire, dont le rang s'obtient directement.

  . **Calcul de l'inverse d'une matrice carrée** (s'il existe)

  Soit $A \in \mathcal{M}_n(\mathbb{K})$. Si A est inversible, il existe P_1, P_2, …, $P_m \in \mathrm{GL}_n(\mathbb{K})$ (matrices élémentaires) telles que : $P_m P_{m-1} \dots P_1 A = \mathrm{I}_n$.

  Dans ce cas, $A^{-1} = P_m P_{m-1} \dots P_1 \mathrm{I}_n$, ce qui signifie que A^{-1} se déduit de I_n par la *même* suite d'opérations sur les lignes que celle utilisée pour déduire I_n de A.

  ✓ On peut aussi utiliser des opérations sur les colonnes, ce qui revient à multiplier à droite par des matrices élémentaires, mais il ne faut pas mélanger les opérations sur les lignes et les colonnes !

[S3.14] Déterminants

- **Applications n–linéaires**

Soit $n \in \mathbb{N}^*$. Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f\colon E^n \to F$ est dite n–linéaire si, pour tout $j \in [\![1\,;n]\!]$ et pour toute famille $(a_i)_{i\in[\![1;n]\!]\setminus\{j\}}$ composée de $(n-1)$ vecteurs de E, l'application partielle :

$$f_j : \begin{cases} E & \longrightarrow & F \\ x & \longmapsto & f(a_1,\dots,a_{j-1},x,a_{j+1},\dots,a_n) \end{cases}$$

est une application linéaire sur E.

Lorsque $F = \mathbb{K}$, on parle de *forme* n-linéaire.

- **Applications n-linéaires symétriques, antisymétriques**

$\mathfrak{S}_n$ désigne le groupe symétrique d'ordre n, c'est-à-dire l'ensemble des permutations de $[\![1\,;n]\!]$ (cf. **[S1.8]**).

Une application n-linéaire de E^n dans F est dite :

. symétrique si :

$$\forall\, \sigma \in \mathfrak{S}_n,\ \forall\, (x_i)_{1\leqslant i\leqslant n} \in E^n,\ f(x_{\sigma(1)},\dots,x_{\sigma(n)}) = f(x_1,\dots,x_n).$$

. antisymétrique si :

$$\forall\, \sigma \in \mathfrak{S}_n,\ \forall\, (x_i)_{1\leqslant i\leqslant n} \in E^n,\ f(x_{\sigma(1)},\dots,x_{\sigma(n)}) = \varepsilon(\sigma) f(x_1,\dots,x_n)$$

(où $\varepsilon(\sigma)$ désigne la signature de la permutation σ).

- **Applications n-linéaires alternées**

  Soit f une application n-linéaire de E^n dans F, $n \geqslant 2$.

  f est dite alternée si, pour toute famille $(x_i)_{1\leqslant i\leqslant n} \in E^n$, on a :

  $$f(x_1,\dots,x_n) = 0 \text{ dès qu'il existe deux indices } i \neq j \text{ tels que } x_i = x_j\,.$$

  Propriétés

  – f antisymétrique $\iff$ f alternée.

  – Si f est une application n-linéaire alternée, alors :

  – $(x_1,\dots,x_n)$ liée $\implies f(x_1,\dots,x_n) = 0$;

  – on ne change pas la valeur de $f(x_1,\dots,x_n)$ si l'on ajoute à un des vecteurs une combinaison linéaire des *autres*.

- **Déterminant d'une famille de vecteurs**

  Soient :

  – E un $\mathbb{K}$-espace vectoriel de dimension n $(n \in \mathbb{N}^*)$;

  – $\mathscr{B} = (e_1,\dots,e_n)$ une base de E ;

  – $(v_j)_{1\leqslant j\leqslant n}$ une famille de vecteurs de E, et $A = (a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{K})$ la matrice de cette famille dans la base $\mathscr{B}$, de sorte que :

  $$\forall\, j \in [\![1\,;n]\!],\ v_j = \sum_{i=1}^{n} a_{ij} e_i\,.$$

  Alors pour toute forme n-linéaire alternée φ sur E^n on a :

  $$\varphi(v_1,\dots,v_n) = \varphi(e_1,\dots,e_n) \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}\,.$$

  Le scalaire : $\det_{\mathscr{B}}(v_1,\dots,v_n) = \displaystyle\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$ s'appelle le déterminant dans la base $\mathscr{B}$ du système de vecteurs $(v_1,\dots,v_n)$.

  ✓ Le déterminant d'une famille de vecteurs dépend de la base $\mathscr{B}$ choisie.

  ✓ Le déterminant n'est défini que pour une famille de vecteurs dont le cardinal est égal à la dimension de E.

  Propriétés

  – Avec les notations précédentes, on a aussi :

  $$\det_{\mathscr{B}}(v_1,\dots,v_n) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}\,.$$

  – L'application $\det_{\mathscr{B}} : E^n \longrightarrow \mathbb{K}$ est une forme n-linéaire alternée.

  – $\det_{\mathscr{B}}(\mathscr{B}) = 1$.

  – Pour toute forme n-linéaire alternée φ sur E, il existe un scalaire λ tel que $\varphi = \lambda \det_{\mathscr{B}}$ (autrement dit, l'ensemble des formes n-linéaires alternées est la droite vectorielle engendrée par $\det_{\mathscr{B}}$).

– $\det_{\mathscr{B}}$ est l'unique forme n-linéaire alternée φ telle que $\varphi(\mathscr{B}) = 1$.

– Pour toute permutation $\sigma \in \mathfrak{S}_n$, on a :

$$\det_{\mathscr{B}}(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma) \det_{\mathscr{B}}(v_1, \dots, v_n)\,.$$

En particulier, si l'on échange deux vecteurs dans la famille $(v_1, \dots, v_n)$, le déterminant change de signe.

– Pour tout $\lambda \in \mathbb{K}$, $\det_{\mathscr{B}}(\lambda v_1, \dots, \lambda v_n) = \lambda^n \det_{\mathscr{B}}(v_1, \dots, v_n)$.

– On ne change pas le déterminant d'un système de vecteurs si l'on ajoute à l'un d'entre eux une combinaison linéaire des *autres*.

- **Changement de base**

Soient $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Alors :

(i) $\det_{\mathscr{B}'}(v_1, \dots, v_n) = \det_{\mathscr{B}'}(\mathscr{B}) \times \det_{\mathscr{B}}(v_1, \dots, v_n)\,.$

(ii) $\det_{\mathscr{B}'}(\mathscr{B}) \times \det_{\mathscr{B}}(\mathscr{B}') = 1\,.$

- **Caractérisation des bases**

Soit $(v_1, \dots, v_n)$ une famille de vecteurs d'un espace vectoriel E de dimension n, et $\mathscr{B}$ une base de E. Alors :

$$(v_1, \dots, v_n) \text{ est une base de } E \iff \det_{\mathscr{B}}(v_1, \dots, v_n) \neq 0\,.$$

- **Déterminant d'un endomorphisme**

E désigne toujours un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$. Soit $u \in \mathscr{L}(E)$.
Le scalaire $\det_{\mathscr{B}}\big(u(\mathscr{B})\big)$ est indépendant de la base $\mathscr{B}$ choisie.
Ce scalaire s'appelle le déterminant de l'endomorphisme u, noté $\det u$.

Propriétés

– Pour tout endomorphisme u de E, pour toute famille $(v_j)_{1 \leqslant j \leqslant n} \in E^n$ et pour toute base $\mathscr{B}$ de E :

$$\det_{\mathscr{B}}\big(u(v_1), \dots, u(v_n)\big) \;=\; \det u \times \det_{\mathscr{B}}(v_1, \dots, v_n)\,.$$

– $\det \mathrm{Id}_E = 1$.

– $\forall\, u \in \mathscr{L}(E),\ \forall\, \lambda \in \mathbb{K},\ \det(\lambda u) = \lambda^n \det u$.

– $\forall\, u \in \mathscr{L}(E),\ u \in \mathrm{GL}(E) \iff \det u \neq 0$.

– $\forall\, u, v \in \mathscr{L}(E),\ \det(u \circ v) = \det u \times \det v$.

– $\forall\, u \in \mathrm{GL}(E),\ \det(u^{-1}) = \dfrac{1}{\det u}\cdot$

– L'ensemble des automorphismes de E de déterminant égal à 1 est un sous-groupe du groupe $\mathrm{GL}(E)$, appelé groupe spécial linéaire, et noté $\mathrm{SL}(E)$.

- **Déterminant d'une matrice carrée**

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$. On appelle déterminant de A le déterminant des vecteurs colonnes de A dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$:

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}\,.$$

Si A est la matrice dans une base $\mathscr{B}$ d'un $\mathbb{K}$-espace vectoriel E d'un endomorphisme u, alors $\det A = \det u$.

Propriétés

- $\det \mathrm{I}_n = 1$.
- $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \forall\, \lambda \in \mathbb{K},\ \det(\lambda A) = \lambda^n \det A$.
- $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ A \in \mathrm{GL}_n(\mathbb{K}) \iff \det A \neq 0$.
- $\forall\, A,\ B \in \mathcal{M}_n(\mathbb{K}),\ \det(AB) = \det A \times \det B = \det(BA)$.
- $\forall\, A \in \mathrm{GL}_n(\mathbb{K}),\ \det(A^{-1}) = \dfrac{1}{\det A}$.
- Deux matrices semblables ont même déterminant.
- $\forall\, A \in \mathcal{M}_n(\mathbb{K})$, $\det({}^{\mathrm{t}}A) = \det A$.

- **Calcul du déterminant par opérations élémentaires**
  - Le déterminant d'une matrice est multiplié par -1 lors de l'échange de deux de ses colonnes ou de deux de ses lignes.
  - Pour tout scalaire λ, le déterminant d'une matrice est multiplié par λ lorsqu'une colonne ou une ligne est multipliée par λ.
  - Le déterminant d'une matrice reste inchangé en ajoutant à l'une de ses colonnes une combinaison linéaire des *autres* colonnes ou en ajoutant à l'une de ses lignes une combinaison linéaire des *autres* lignes.

- **Calcul d'un déterminant par blocs**

  Soit $A \in \mathcal{M}_n(\mathbb{K})$, écrite par blocs sous la forme $A = \begin{bmatrix} A_1 & B_2 \\ 0 & A_2 \end{bmatrix}$, où A_1 et A_2 sont deux blocs *carrés*. Alors $\det A = \det A_1 \times \det A_2$.

  En particulier, le déterminant d'une matrice triangulaire est égal au produit de ses termes diagonaux.

- **Développement suivant une rangée**
  - Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$. Pour tout $(i,j) \in [\![1\,;n]\!]^2$, on note Δ_{ij} le déterminant obtenu en supprimant la i-ème ligne et la j-ème colonne de A. On dit alors que Δ_{ij} est le mineur d'indice (i,j) de la matrice A.
  - Développement de $\det A$ selon la j-ème colonne :

$$\forall\, j \in [\![1\,;n]\!],\ \det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}\ .$$

  - Développement de $\det A$ selon la i-ème ligne :

$$\forall\, i \in [\![1\,;n]\!],\ \det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij}\ .$$

  - Le terme $(-1)^{i+j}\Delta_{ij}$ s'appelle le cofacteur d'indice (i,j) de la matrice A.

- **Comatrice**

  Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle comatrice de A, notée $\mathrm{com}(A)$, la matrice formée des cofacteurs de A, soit plus précisément :

$$\forall\, (i,j) \in [\![1\,;n]\!]^2,\ \big(\mathrm{com}(A)\big)_{ij} = (-1)^{i+j}\Delta_{ij} \text{ avec } \Delta_{ij} \text{ mineur d'indice } (i,j)\,.$$

  On appelle matrice complémentaire de A la matrice $\widetilde{A} = {}^{\mathrm{t}}\mathrm{com}(A)$. On a alors :

- $\forall\, A \in \mathcal{M}_n(\mathbb{K})$, $A\widetilde{A} = \widetilde{A}A = \det A.\mathrm{I}_n$.
- Si $A \in \mathrm{GL}_n(\mathbb{K})$, alors : $A^{-1} = \dfrac{1}{\det A}\widetilde{A}$.

- **Déterminant de Vandermonde**

  On appelle déterminant de Vandermonde un déterminant de la forme :

$$V(a_1,a_2,\ldots,a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \ldots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{vmatrix},$$

  où $(a_1,\ldots,a_n) \in \mathbb{K}^n$.

  On sait calculer ce déterminant :

$$V(a_1,a_2,\ldots,a_n) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i).$$

  Il en résulte que :

$$V(a_1,a_2,\ldots,a_n) \neq 0 \iff \text{les } a_i \text{ sont deux à deux distincts}.$$

[S3.15] Systèmes linéaires

- Un système linéaire de p équations à q inconnues est un système de la forme :

$$(S) \quad \begin{cases} a_{11}x_1 + \ldots + a_{1q}x_q &= b_1 \\ a_{21}x_1 + \ldots + a_{2q}x_q &= b_2 \\ \vdots & \vdots \\ a_{p1}x_1 + \ldots + a_{pq}x_q &= b_p \end{cases}$$

  où les coefficients $(a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}}$ et $(b_i)_{1 \leqslant i \leqslant p}$ sont donnés dans $\mathbb{K}$. $x_1,\ldots,x_q$, éléments de $\mathbb{K}$, sont les inconnues.

  Le système est dit homogène si $b_1 = \ldots = b_p = 0$.

- **Interprétations**

  Un système linéaire peut être interprété de différentes façons, chacune correspondant à un problème donné.

  - La matrice $A = (a_{ij})_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} \in \mathcal{M}_{p,q}(\mathbb{K})$ s'appelle la matrice du système ; le rang de A s'appelle le rang du système.

    Soit alors $B = {}^{\mathrm{t}}\begin{pmatrix} b_1 & \ldots & b_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$ et $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \ldots & x_q \end{pmatrix} \in \mathcal{M}_{q,1}(\mathbb{K})$; le système (S) s'écrit alors matriciellement sous la forme

$$AX = B.$$

  - En notant $C_1,\ldots,C_q$ les vecteurs colonnes de A, on peut aussi écrire le système sous la forme : $\displaystyle\sum_{j=1}^{q} x_j C_j = B$.

Résoudre le système équivaut donc à écrire B comme combinaison linéaire des C_j.
On notera que le rang du système est aussi égal au rang du système de vecteurs $(C_1,\dots,C_q)$.

- Soient E et F deux $\mathbb{K}$-espaces vectoriels tels que $\dim E = q$ et $\dim F = p$. Soient alors $\mathscr{B}_E$ et $\mathscr{B}_F$ deux bases respectivement de E et de F, et $u \in \mathscr{L}(E,F)$ telle que $A = M_{\mathscr{B}_E}^{\mathscr{B}_F}(u)$.
Le système (S) peut alors s'écrire $u(x) = b$, avec $x \in E$ de coordonnées $(x_1,\dots,x_q)$ dans $\mathscr{B}_E$ et $b \in F$ de coordonnées $(b_1,\dots,b_p)$ dans $\mathscr{B}_F$.
Résoudre le système revient ici à trouver les antécédents de b par u.
On notera que le rang du système est aussi égal au rang de l'application linéaire u.
- Avec les notations précédentes, en notant $\varphi_1,\dots,\varphi_p$ les formes linéaires sur E de matrices respectives les vecteurs lignes de A (c'est-à-dire $\varphi_i : x \mapsto \sum_{j=1}^{q} a_{ij}x_j$),
le système (L) s'écrit : $\left\{\begin{array}{ccc} \varphi_1(x) & = & b_1 \\ \vdots & & \vdots \\ \varphi_p(x) & = & b_p \end{array}\right.$.
Résoudre le système revient ici à chercher l'intersection de p hyperplans affines.
On notera que le rang du système est aussi égal au rang de la famille $(\varphi_1,\dots,\varphi_p)$ dans l'espace dual E^* (car $\operatorname{rg} A = \operatorname{rg} {}^tA$).
Cas particulier : dans le cas d'un système homogène, qui s'écrit donc
$$\forall\, i \in [\![1\,;p]\!],\ \varphi_i(x) = 0\,,$$
l'ensemble des solutions est alors égal à $\bigcap_{i=1}^{p} \operatorname{Ker}\varphi_i$; c'est l'intersection de p hyperplans de E, et c'est un espace de dimension $q - r$, où r est le rang du système (d'après le théorème du rang, puisque cet espace est aussi égal à $\operatorname{Ker} u$).

- **Systèmes de Cramer**

Un système linéaire (S) est dit de Cramer si sa matrice associée est inversible.
Tout système de Cramer admet une solution unique.

Soit un système de Cramer de taille n écrit sous la forme : $\sum_{j=1}^{n} x_j C_j = B$.

On a alors, pour tout $i \in [\![1\,;n]\!]$:

$$\begin{aligned}\det(C_1,\dots,C_{i-1},B,C_{i+1},\dots,C_n) &= \sum_{j=1}^{n} x_j \det(C_1,\dots,C_{i-1},C_j,C_{i+1},\dots,C_n)\\ &= x_i \det(C_1,\dots,C_{i-1},C_i,C_{i+1},\dots,C_n)\\ &= x_i \det A\,.\end{aligned}$$

On a donc les *formules de Cramer* :
$$\forall\, i \in [\![1\,;n]\!],\ x_i = \frac{\det A_i(B)}{\det A}\,,$$
où $A_i(B)$ est la matrice obtenue en remplaçant la i^{e} colonne de A par B.

- **Structure de l'ensemble des solutions**

  On note $\mathscr{S}$ l'ensemble des solutions de (S). On dit que (S) est compatible si $\mathscr{S} \neq \emptyset$, et incompatible sinon.

  L'ensemble des solutions du système homogène associé à (S) est $\operatorname{Ker} A = \{X \in \mathcal{M}_{q,1}(\mathbb{K}) \mid AX = 0\}$. Si (S) est compatible, alors il existe $X_0 \in \mathcal{M}_{p,q}(\mathbb{K})$ tel que $AX_0 = B$; X_0 s'appelle une solution particulière du système. L'ensemble des solutions du système est alors le sous-espace affine passant par X_0 et de direction $\operatorname{Ker} A$:

$$\mathscr{S} = X_0 + \operatorname{Ker} A = \{X_0 + X \mid X \in \operatorname{Ker} A\} .$$

- **Résolution dans le cas général**

  . **Résolution avec les formules de Cramer**

  Soit un système de p équations à q inconnues $AX = B$, et r le rang de ce système ($r \leqslant p$ et $r \leqslant q$).

  On appelle matrice principale du système toute matrice carrée inversible d'ordre r extraite de A (on sait qu'il en existe).

  Si A_r est la matrice principale choisie, les inconnues dont les indices sont ceux des colonnes correspondantes de A_r sont appelées inconnues principales, et les équations dont les indices sont ceux des lignes de A_r sont appelées équations principales.

  Le système est compatible si et seulement si les $p - r$ équations secondaires sont combinaisons linéaires des r équations principales ; dans ce cas, le système équivaut à celui formé par ces r équations principales ; en résolvant un système de Cramer de matrice A_r, on peut alors exprimer les r inconnues principales en fonction des $q - r$ inconnues secondaires.

  . **Résolution par la méthode du pivot**

  Pour résoudre le système $AX = B$, on peut utiliser des opérations élémentaires sur les lignes de la matrice complète du système $\begin{bmatrix} A & B \end{bmatrix}$ pour obtenir un système équivalent dont la matrice associée A' est échelonnée par lignes.

  Les inconnues principales sont alors celles correspondant aux pivots de A'. Le système sera compatible si et seulement si les $q - r$ dernières équations s'écrivent « $0 = 0$ ». Dans ce cas, le nouveau système obtenu permettra d'exprimer facilement les inconnues principales en fonction des inconnues secondaires.

  > ✓ Le choix des inconnues principales et secondaires est souvent arbitraire. Cependant leur nombre est toujours le même : si r est le rang du système, il y a r inconnues principales et $q - r = \dim(\operatorname{Ker} A)$ inconnues secondaires.

Thème 4 - Réduction des endomorphismes

Réduire un endomorphisme $u \in \mathscr{L}(E)$ consiste, dans le cas le plus général, à trouver une famille $(E_i)_{1\leqslant i\leqslant p}$ de sous-espaces vectoriels de E, non réduits à $\{0\}$, stables par u, tels que $E = \bigoplus_{i=1}^{p} E_i$.

En dimension finie, cela revient à trouver une base $\mathscr{B}$ de E où la matrice de u est diagonale par blocs :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} \boxed{A_1} & & & 0 \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_p} \end{pmatrix}$$

chaque bloc A_i étant la matrice dans une base de E_i de l'endomorphisme u_i induit par u sur E_i (cf. **[S3.8]**).

Le cas le plus simple est celui où les E_i sont de dimension 1, ce qui revient à chercher les droites vectorielles stables par u.

[S4.1] Éléments propres d'un endomorphisme

- Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E.
  - On dit que $\lambda \in \mathbb{K}$ est une valeur propre de u s'il existe un vecteur $x \in E$, $x \neq 0_E$, tel que $u(x) = \lambda x$.
  - On dit que $x \in E$ est un vecteur propre de u si $x \neq 0_E$ et s'il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.
    On dit alors que x est *un* vecteur propre associé à la valeur propre λ, ou que λ est *la* valeur propre associé au vecteur propre x.
    Cela équivaut à dire que la droite vectorielle $\mathbb{K}.x$ est stable par u.
  - On appelle spectre de u, noté $\mathrm{Sp}_{\mathbb{K}}(u)$ ou plus simplement $\mathrm{Sp}(u)$, l'ensemble des valeurs propres de u (dans $\mathbb{K}$).
- Il est très important de bien connaître et comprendre les équivalences suivantes :

$$\begin{aligned} \lambda \in \mathrm{Sp}(u) &\iff \exists\, x \neq 0 \text{ tel que } u(x) = \lambda x \\ &\iff \exists\, x \neq 0 \text{ tel que } (u - \lambda \mathrm{Id}_E)(x) = 0 \\ &\iff \mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0\} \\ &\iff (u - \lambda \mathrm{Id}_E) \ \textit{non injective.} \end{aligned}$$

  En particulier : 0 est valeur propre de $u \iff u$ non injectif.
- Si $\lambda \in \mathrm{Sp}(u)$, on appelle sous-espace propre de u associé à la valeur propre λ le sous-espace vectoriel :

$$E_\lambda(u) = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)\,.$$

Ainsi, $E_\lambda(u) = \{x \in E \mid u(x) = \lambda x\}$ est constitué des vecteurs propres de u associés à la valeur propre λ et du vecteur nul.

✓ On rappelle que le vecteur nul n'est *pas* un vecteur propre !

- La propriété suivante, qui découle de la définition, est très souvent utilisée.

  Le sous-espace propre associé à la valeur propre 0 est le noyau de u.

On a aussi une propriété importante concernant les valeurs propres non nulles.

Les sous-espaces propres de u associés à une valeur propre non nulle sont inclus dans $\operatorname{Im} u$.

(En effet, si x est un vecteur propre associé à la valeur propre $\lambda \neq 0$ on a $u(x) = \lambda x$ donc $x = u\left(\frac{x}{\lambda}\right) \in \operatorname{Im} u$).

[S4.2] Propriétés des sous-espaces propres

u désigne toujours ici un endomorphisme d'un espace vectoriel E.

- Si λ est une valeur propre de u, $E_\lambda(u)$ est stable par u, et l'endomorphisme induit par u sur $E_\lambda(u)$ est l'homothétie de rapport λ.
- Si $v \in \mathscr{L}(E)$ *commute* avec u, $E_\lambda(u)$ est stable par v.
- Soit $p \in \mathbb{N}^*$ et $(\lambda_i)_{1 \leqslant i \leqslant p}$ une famille de valeurs propres de u *deux à deux distinctes.*

  Alors les sous-espaces propres $E_{\lambda_i}(u)$ pour $1 \leqslant i \leqslant p$ sont en somme directe.
- Soit $(\lambda_i)_{1 \leqslant i \leqslant p}$ une famille de valeurs propres de u *deux à deux distinctes* et, pour tout $i \in [\![1\,;p]\!]$, x_i un vecteur propre associé à la valeur propre λ_i.

  Alors la famille $(x_i)_{1 \leqslant i \leqslant p}$ est *libre.*
- On en déduit immédiatement que, dans un espace vectoriel de dimension n, un endomorphisme possède au plus n valeurs propres distinctes.

[S4.3] Éléments propres d'une matrice

- Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$.
  - On dit que $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe un vecteur colonne $V \in \mathcal{M}_{n,1}(\mathbb{K})$, $V \neq 0$, tel que $AV = \lambda V$.
  - On dit alors que le vecteur colonne $V \in \mathcal{M}_{n,1}(\mathbb{K})$ est un vecteur propre de A associé à la valeur propre λ.
  - On appelle spectre de A, noté $\operatorname{Sp}_{\mathbb{K}}(A)$ ou plus simplement $\operatorname{Sp}(A)$, l'ensemble des valeurs propres de A (dans $\mathbb{K}$).
  - Si $\lambda \in \operatorname{Sp}(A)$, on appelle sous-espace propre de A associé à λ l'ensemble $E_\lambda(A) = \{V \in \mathcal{M}_{n,1}(\mathbb{K}) \mid AV = \lambda V\}$. C'est évidemment un sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbb{K})$.

    ✓ Lorsque A est à coefficients réels, il est important de distinguer ses valeurs propres réelles et ses valeurs propres complexes, c'est-à-dire les ensembles $\operatorname{Sp}_{\mathbb{R}}(A)$ et $\operatorname{Sp}_{\mathbb{C}}(A)$.Dans le premier cas, les vecteurs propres seront des

vecteurs colonnes à chercher dans $\mathcal{M}_{n,1}(\mathbb{R})$ et dans le second cas dans $\mathcal{M}_{n,1}(\mathbb{C})$.

De plus, si λ est une valeur propre complexe de la matrice A à coefficients réels, alors $\overline{\lambda}$ est encore une valeur propre de A et

$$V \in E_\lambda(A) \iff \overline{V} \in E_{\overline{\lambda}}(A)\,.$$

- **Lien entre éléments propres d'une matrice et d'un endomorphisme**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$ et $\mathscr{B}$ une base de E.
  Soit u un endomorphisme de E et A la matrice de u dans la base $\mathscr{B}$.
  - $\lambda \in \mathbb{K}$ est valeur propre de $A \iff \lambda$ est valeur propre de u.
  - Soit $x \in E$ et V la matrice colonne de ses coordonnées dans $\mathscr{B}$.

$$x \in E_\lambda(u) \iff V \in E_\lambda(A)\,.$$

[S4.4] Polynôme caractéristique

E désigne ici un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^*$.

- Soit $u \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$. Par définition d'une valeur propre on a :

$$\lambda \in \mathrm{Sp}(u) \iff \det(\lambda \mathrm{Id}_E - u) = 0\,.$$

  Soit $\mathscr{B}$ une base de E, $A = (a_{ij})$ la matrice de u dans $\mathscr{B}$. Alors :

$$\det(\lambda \mathrm{Id}_E - u) = \det(\lambda \mathrm{I}_n - A) = \begin{vmatrix} \lambda - a_{11} & \dots & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ -a_{n1} & \dots & \dots & \lambda - a_{nn} \end{vmatrix}.$$

  D'après la formule donnant la valeur du déterminant d'une matrice, il apparaît clairement que $\det(\lambda \mathrm{Id}_E - u)$ est une fonction polynôme en λ, de degré n.
  On appelle polynôme caractéristique de u le polynôme, noté χ_u, défini par :

$$\forall \lambda \in \mathbb{K},\ \chi_u(\lambda) = \det(\lambda \mathrm{Id}_E - u) \quad \text{ou} \quad \chi_u = \det(X \mathrm{Id}_E - u)\,.$$

  Si $\mathscr{B}$ est une base quelconque de E, et si $A = M_{\mathscr{B}}(u)$, on a aussi :

$$\forall \lambda \in \mathbb{K},\ \chi_u(\lambda) = \det(\lambda \mathrm{I}_n - A) \quad \text{ou} \quad \chi_u = \det(X \mathrm{I}_n - A)\,.$$

  Le polynôme $\chi_A = \det(X \mathrm{I}_n - A)$ s'appelle naturellement le polynôme caractéristique de la matrice A.

- Les valeurs propres d'un endomorphisme de E (ou d'une matrice carrée) sont exactement les racines dans $\mathbb{K}$ de son polynôme caractéristique.

  ✓ Comme il a déjà été dit, lorsque A est une matrice à coefficients réels, il convient de distinguer ses valeurs propres dans $\mathbb{R}$ ou dans $\mathbb{C}$, c'est-à-dire les racines de χ_A dans $\mathbb{R}$ ou dans $\mathbb{C}$.

- Deux matrices semblables ont même polynôme caractéristique (donc mêmes valeurs propres).

- Le polynôme caractéristique d'une matrice est égal à celui de sa transposée. Donc pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, $\operatorname{Sp} A = \operatorname{Sp} {}^tA$.
- Le polynôme caractéristique de u est un polynôme unitaire de degré n. Plus précisément, il s'écrit :

$$\chi_u = X^n - (\operatorname{tr} u)X^{n-1} + \cdots + (-1)^n \det u\,.$$

- Si le polynôme caractéristique χ_u de u est *scindé* (ce qui est toujours le cas si $\mathbb{K} = \mathbb{C}$), on peut écrire :

$$\chi_u = \prod_{i=1}^{n}(X - \lambda_i)\,,$$

où les λ_i sont les valeurs propres, distinctes ou non, de u. À l'aide des relations coefficients-racines on obtient alors les relations (*très importantes*) :

$$\operatorname{tr} u = \lambda_1 + \cdots + \lambda_n \quad ; \quad \det u = \lambda_1\lambda_2\cdots\lambda_n\,.$$

✓ On prendra soin cependant de n'utiliser ces relations que lorsque le polynôme caractéristique de u est *scindé*. Ainsi, dans le cas d'une matrice à coefficients réels, il faut considérer toutes les valeurs propres dans $\mathbb{C}$.

- **Ordre de multiplicité d'une valeur propre**

  Soit $u \in \mathscr{L}(E)$ et soit $\lambda \in \operatorname{Sp}_{\mathbb{K}}(u)$.

  On dit que λ est valeur propre d'ordre de multiplicité m_λ de u si λ est une racine d'ordre m_λ du polynôme caractéristique χ_u de u.

  On a alors :

$$1 \leqslant \dim E_\lambda(u) \leqslant m_\lambda\,.$$

[S4.5] Endomorphismes ou matrices diagonalisables

- Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel E de dimension finie est dit diagonalisable s'il existe une base de E formée de vecteurs propres de u.

  Cela équivaut à dire qu'il existe une base $\mathscr{B}$ de E dans laquelle la matrice de u est diagonale ; $\mathscr{B}$ est formée de vecteurs propres de u, et les valeurs propres de u (éventuellement confondues) sont les éléments de la diagonale de $M_{\mathscr{B}}(u)$.

- **Diagonalisabilité et sous-espaces propres**

  Un endomorphisme u est diagonalisable si et seulement si il vérifie l'une des propriétés (équivalentes) suivantes :

  (i) E est somme (directe) des sous-espaces propres de u : $E = \bigoplus_{\lambda \in \operatorname{Sp}(u)} E_\lambda(u)$,

  (ii) $\dim E = \sum_{\lambda \in \operatorname{Sp}(u)} \dim\left(E_\lambda(u)\right)$,

  (iii) χ_u est scindé dans $\mathbb{K}[X]$ et, pour tout $\lambda \in \operatorname{Sp}(u)$, on a : $m_\lambda = \dim E_\lambda(u)$ (en notant m_λ l'ordre de multiplicité de la valeur propre λ).

Cas particulier : si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension n qui possède n valeurs propres distinctes, alors u est diagonalisable (et, dans ce cas, tous ses sous-espaces propres sont de dimension 1).

✓ Attention ! il ne s'agit là que d'une condition *suffisante*, et non nécessaire, de diagonalisabilité.
Considérer par exemple $u = \mathrm{Id}_E$: u est diagonalisable, mais u admet une seule valeur propre, 1.

[S4.6] Polynômes d'endomorphismes

- Soit E un $\mathbb{K}$-espace vectoriel, et $u \in \mathscr{L}(E)$.

  À tout polynôme $P = \sum_{k=0}^{d^\circ(P)} a_k X^k$ de $\mathbb{K}[X]$ on peut associer l'endomorphisme de E : $P(u) = \sum_{k=0}^{d^\circ(P)} a_k u^k$ (où u^k désigne, comme d'habitude, l'endomorphisme défini par : $u^0 = \mathrm{Id}_E$ et, pour tout $k \geqslant 1$: $u^k = u \circ u^{k-1}$).

  ✓ Si x est un vecteur de E, l'écriture $P(u)(x)$ a un sens (c'est l'image de x par l'endomorphisme $P(u)$) mais l'écriture $P(u(x))$ ne veut rien dire !

- u étant un endomorphisme donné de E, l'application $\varphi_u : \begin{cases} \mathbb{K}[X] & \longrightarrow & \mathscr{L}(E) \\ P & \longmapsto & P(u) \end{cases}$ est un morphisme de $\mathbb{K}$-algèbres, c'est-à-dire que, pour P, $Q \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$:
  - si P est le polynôme constant égal à 1, $P(u) = \mathrm{Id}_E$;
  - $(P+Q)(u) = P(u) + Q(u)$;
  - $(\lambda P)(u) = \lambda P(u)$;
  - $(PQ)(u) = P(u) \circ Q(u)$.

  On note $\mathbb{K}[u] = \mathrm{Im}\,\varphi_u = \{P(u) \mid P \in \mathbb{K}[X]\}$ l'algèbre des polynômes de l'endomorphisme u ; c'est une sous-algèbre commutative de $\mathscr{L}(E)$: si P, $Q \in \mathbb{K}[X]$, les endomorphismes $P(u)$ et $Q(u)$ commutent.

- Soit E un $\mathbb{K}$-espace vectoriel, et u, $v \in \mathscr{L}(E)$, tels que $u \circ v = v \circ u$.
  Alors pour tous P, $Q \in \mathbb{K}[X]$, $\mathrm{Im}\big(P(u)\big)$ et $\mathrm{Ker}\big(P(u)\big)$ sont stables par $Q(v)$.

- Soit F un sous-espace vectoriel de E, stable par $u \in \mathscr{L}(E)$ et soit $P \in \mathbb{K}[X]$.
  Alors F est stable par $P(u)$.

- Si λ est une valeur propre de u et si $P \in \mathbb{K}[X]$, alors $P(\lambda)$ est une valeur propre de $P(u)$.

[S4.7] Polynômes annulateurs, polynôme minimal

- Soit $u \in \mathscr{L}(E)$. Un polynôme $P \in \mathbb{K}[X]$ s'appelle un polynôme annulateur de u si $P(u) = 0$ (endomorphisme nul).

L'ensemble des polynômes annulateurs de u est donc le noyau du morphisme d'algèbres $\varphi_u : \begin{cases} \mathbb{K}[X] & \longrightarrow & \mathscr{L}(E) \\ P & \longmapsto & P(u) \end{cases}$; c'est un idéal de $\mathbb{K}[X]$, appelé idéal annulateur de u et noté $\mathrm{Ann}(u)$.

Puisque l'anneau $\mathbb{K}[X]$ est principal,

- soit $\mathrm{Ann}(u) = \{0\}$ (et alors φ_u est injective) ;
- soit il existe un et un seul polynôme unitaire (non nul), noté π_u, tel que $\mathrm{Ann}(u)$ soit exactement l'ensemble des polynômes multiples de π_u.

  π_u s'appelle alors le polynôme minimal de u.

Le polynôme minimal d'un endomorphisme u (s'il existe) est donc caractérisé par :

$$\begin{cases} \pi_u \neq 0 \text{ et } \pi_u \text{ unitaire} ; \\ \pi_u(u) = 0_{\mathscr{L}(E)} ; \\ \forall\, P \in \mathbb{K}[X],\ P(u) = 0_{\mathscr{L}(E)} \implies P \text{ multiple de } \pi_u\,. \end{cases}$$

- Si E est un $\mathbb{K}$-espace vectoriel *de dimension finie*, tout endomorphisme $u \in \mathscr{L}(E)$ possède un polynôme minimal.

  ✓ Lorsque E n'est pas de dimension finie, un endomorphisme peut ne pas posséder de polynôme annulateur autre que le polynôme nul (considérer par exemple l'endomorphisme u de $\mathbb{K}[X]$ qui à tout polynôme P associe son polynôme dérivé P'.)

- **Deux exemples importants d'utilisation d'un polynôme annulateur**

  Soit $u \in \mathscr{L}(E)$, possédant un polynôme annulateur (non nul) P.

  . **Calcul de l'inverse**

  Si P s'écrit $P = \sum_{k=0}^{n} a_k X^k$ avec $\underline{a_0 \neq 0}$, la relation $P(u) = 0$ s'écrit aussi : $u \circ \left(\sum_{k=1}^{n} a_i u^{i-1} \right) = -a_0 \mathrm{Id}_E$, ce qui prouve que u est inversible et permet d'exprimer u^{-1} à l'aide des puissances de u.

  . **Calcul des puissances**

  Pour tout $n \in \mathbb{N}$, la division euclidienne de X^n par P s'écrit : $X^n = PQ_n + R_n$ avec $\deg(R_n) < \deg(P)$.

  On a alors : $u^n = P(u) \circ Q_n(u) + R_n(u) = R_n(u)$, ce qui permet d'exprimer u^n à l'aide des premières puissances de u.

[S4.8] Polynômes de matrices

- Soit $M \in \mathcal{M}_n(\mathbb{K})$. À tout polynôme $P = \sum_{k=0}^{d^\circ(P)} a_k X^k$ de $\mathbb{K}[X]$, on peut associer la matrice $P(M) = \sum_{k=0}^{d^\circ(P)} a_k M^k$ (avec $M^0 = I_n$).

Il est clair que, si $M = M_{\mathscr{B}}(u)$, où u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension n rapporté à une base $\mathscr{B}$, on aura $P(M) = M_{\mathscr{B}}\big(P(u)\big)$.

On dira qu'un polynôme P est annulateur de M si $P(M) = 0$. Cela équivaut à dire que P est annulateur de u.

Puisque u, endomorphisme d'un espace vectoriel de dimension finie, possède un polynôme annulateur non nul, on en déduit que M possède un polynôme annulateur non nul. On peut donc définir le polynôme minimal π_M de M, qui est le même que celui de u.

- Si M est une matrice à coefficients réels, son polynôme minimal en tant que matrice de $\mathcal{M}_n(\mathbb{R})$ est le même que son polynôme minimal en tant que matrice de $\mathcal{M}_n(\mathbb{C})$.
- Si deux matrices carrées A et A' sont semblables et si P est un polynôme annulateur de A, P est aussi un polynôme annulateur de A'.

  En particulier, A et A' ont même polynôme minimal.
- Si $A \in \mathcal{M}_n(\mathbb{K})$ et si P est un polynôme annulateur de A, P est aussi un polynôme annulateur de ${}^tA$.

  En particulier, A et ${}^tA$ ont même polynôme minimal.

[S4.9] Théorème de décomposition des noyaux

La plupart d'entre vous ne connaissent ce théorème que sous sa forme finale (la plus utilisée il est vrai), mais pour l'obtenir, il existe certains théorèmes intermédiaires non dénués d'intérêt propre que nous allons citer.

- Soit E un $\mathbb{K}$-espace vectoriel, u un endomorphisme de E, et P, Q deux polynômes de $\mathbb{K}[X]$.

  Si l'on note $D = \text{pgcd}(P,Q)$ et $M = \text{ppcm}(P,Q)$, on a :

  (i) $\operatorname{Ker} M(u) = \operatorname{Ker} P(u) + \operatorname{Ker} Q(u)$,

  (ii) $\operatorname{Ker} D(u) = \operatorname{Ker} P(u) \cap \operatorname{Ker} Q(u)$.
- Si P et Q sont deux polynômes de $\mathbb{K}[X]$ *premiers entre eux*, alors :

$$\operatorname{Ker}(PQ)(u) = \operatorname{Ker} P(u) \oplus \operatorname{Ker} Q(u)\,.$$

- Si $P_1, \ldots, P_n$ sont n polynômes de $\mathbb{K}[X]$ *premiers entre eux deux à deux*, alors :

$$\operatorname{Ker}\left(\prod_{k=1}^{n} P_k\right)(u) = \bigoplus_{k=1}^{n} \operatorname{Ker} P_k(u)\,.$$

- Si P est un *polynôme annulateur* de u, et si $P = \lambda \prod_{k=1}^{n} P_k^{m_k}$ est sa décomposition en facteurs irréductibles de $\mathbb{K}[X]$ (P_k polynômes irréductibles unitaires et distincts deux à deux, m_k entiers naturels non nuls), alors :

$$E = \bigoplus_{k=1}^{n} \operatorname{Ker} P_k^{m_k}(u)\,.$$

[S4.10] Polynômes annulateurs et réduction

- Soit $u \in \mathscr{L}(E)$ et P un polynôme annulateur non nul de u. Alors toute valeur propre (dans $\mathbb{K}$) de u est une racine de P (dans $\mathbb{K}$).
  Autrement dit : $\mathrm{Sp}(u)$ est *inclus* dans l'ensemble des racines de P.

  ✓ Attention : si P est un polynôme annulateur quelconque de u, toutes ses racines ne sont pas nécessairement valeurs propres de u !

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathscr{L}(E)$.
  Le spectre de u dans $\mathbb{K}$ est exactement l'ensemble des racines dans $\mathbb{K}$ du polynôme minimal π_u.

- Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si il possède un polynôme annulateur scindé dans $\mathbb{K}[X]$ et à racines simples.

- Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si son polynôme minimal est scindé dans $\mathbb{K}[X]$ et à racines simples.

- **Théorème de Cayley-Hamilton**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et u un endomorphisme de E. Alors : $\chi_u(u) = 0_{\mathscr{L}(E)}$.
  Autrement dit, le polynôme caractéristique de u est aussi un polynôme annulateur de u, ou : le polynôme minimal de u divise son polynôme caractéristique.

[S4.11] Sous-espaces stables

- Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie E.
  Soit F un sous-espace vectoriel de E *stable* par u, et $v \in \mathscr{L}(F)$ l'endomorphisme de F induit par u.
  - Le polynôme caractéristique de v divise le polynôme caractéristique de u. En particulier on a :

$$\mathrm{Sp}(v) \subset \mathrm{Sp}(u) \quad \text{et} \quad \forall\, \lambda \in \mathrm{Sp}(v),\ E_\lambda(v) = E_\lambda(u) \cap F\,.$$

  - Le polynôme minimal de v divise le polynôme minimal de u.

- **Diagonalisabilité et sous-espaces stables**
  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathscr{L}(E)$.
  . Si F est un sous-espace vectoriel de E *stable* par u, et si $v \in \mathscr{L}(F)$ est l'endomorphisme de F induit par u, alors :

$$u \text{ diagonalisable} \implies v \text{ diagonalisable}\,.$$

  . *Lorsque u est diagonalisable*, un sous-espace vectoriel F de E est stable par u si et seulement si F admet une base formée de vecteurs propres de u.

✓ Ce résultat tombe en défaut lorsque u n'est pas diagonalisable : par exemple, si u est une rotation d'un espace euclidien de dimension 3, différente de $\pm\mathrm{Id}_E$, le plan orthogonal à l'axe de cette rotation est stable par u mais ne contient aucun vecteur propre de u.

- Dans le cas d'un endomorphisme non diagonalisable mais dont le polynôme caractéristique est scindé, voir **[S4.14]**.

[S4.12] Trigonalisation

E désigne ici un $\mathbb{K}$-espace vectoriel de dimension finie non nulle.

- Un endomorphisme $u \in \mathscr{L}(E)$ est dit trigonalisable s'il existe une base $\mathscr{B}$ de E telle que $M_{\mathscr{B}}(u)$ soit triangulaire supérieure.
  Une matrice carrée $A \in \mathcal{M}(\mathbb{K})$ est dite trigonalisable si l'endomorphisme a de $\mathbb{K}^n$ qui lui est canoniquement associé est trigonalisable.
  Cela revient à dire que A est semblable à une matrice triangulaire supérieure, c'est-à-dire qu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $T \in \mathcal{T}_n^+(\mathbb{K})$ telles que : $T = P^{-1}AP$ (ou $A = PTP^{-1}$).
- Un endomorphisme $u \in \mathscr{L}(E)$ est trigonalisable si et seulement si il possède un polynôme annulateur scindé dans $\mathbb{K}[X]$ (ce qui équivaut à dire que χ_u est scindé dans $\mathbb{K}[X]$).
  En particulier, si $\mathbb{K} = \mathbb{C}$, tout endomorphisme de E est trigonalisable, et toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable.

  ✓ Si $u \in \mathscr{L}(E)$ est trigonalisable et si $\mathscr{B}$ est une base de E dans laquelle la matrice T de u est triangulaire supérieure, alors les éléments diagonaux de T sont exactement les valeurs propres de u (chacune étant comptée avec son ordre de multiplicité).

[S4.13] Endomorphismes nilpotents

- Un endomorphisme u d'un espace vectoriel E est dit nilpotent s'il existe un entier naturel n tel que $u^n = 0$.
  On appelle alors indice de nilpotence de u le plus petit entier p tel que $u^p = 0$.
- Si E est de dimension finie, et si u est nilpotent d'indice p, alors $p \leqslant \dim E$.
- u est un endomorphisme nilpotent d'indice p si et seulement si son polynôme minimal est $\pi_u = X^p$.
- Si u est un endomorphisme d'un espace vectoriel de dimension n, u est nilpotent si et seulement si son polynôme caractéristique est égal à X^n.
- Un endomorphisme u d'un $\mathbb{C}$-espace vectoriel de dimension finie est nilpotent si et seulement si sa seule valeur propre est 0.

  ✓ Ce résultat tombe en défaut si E est un $\mathbb{R}$-espace vectoriel ; par exemple, la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ admet pour seule valeur propre *réelle* 0 mais n'est pas nilpotente.

- Si u est un endomorphisme d'un espace vectoriel de dimension finie, u est nilpotent si et seulement si il existe une base de E dans laquelle la matrice de u est triangulaire supérieure à éléments diagonaux nuls.

[S4.14] Endomorphismes à polynôme annulateur scindé

- Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$ et $u \in \mathscr{L}(E)$. Les propositions suivantes sont équivalentes :

  (i) il existe un polynôme annulateur de u scindé dans $\mathbb{K}[X]$,

  (ii) le polynôme minimal de u est scindé dans $\mathbb{K}[X]$,

  (iii) le polynôme caractéristique de u est scindé dans $\mathbb{K}[X]$.

- **Sous-espaces caractéristiques**

  Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$ et $u \in \mathscr{L}(E)$ annulant un polynôme scindé.

  Son polynôme caractéristique s'écrit alors $\chi_u = \prod_{i=1}^{p}(X-\lambda_i)^{m_i}$, où les $\lambda_i \in \mathbb{K}$ sont les valeurs propres distinctes de u et m_i leurs ordres de multiplicité.
  Pour $i \in [\![1\,;p]\!]$, soit $F_{\lambda_i} = \operatorname{Ker}\big((u-\lambda_i \mathrm{Id}_E)^{m_i}\big)$: ce sont les sous-espaces caractéristiques de u. Alors :

  – $E = \bigoplus_{i=1}^{p} F_{\lambda_i}$;

  – $\forall\, i \in [\![1\,;p]\!],\ \dim F_{\lambda_i} = m_i$;

  – les F_{λ_i} sont stables par u ; si l'on note u_i l'endomorphisme induit par u sur F_{λ_i}, on a les propriétés suivantes :

    ▪ le polynôme minimal π_{u_i} de u_i est égal à $(X-\lambda_i)^{\alpha_i}$, où $1 \leqslant \alpha_i \leqslant m_i$; ainsi, $F_{\lambda_i} = \operatorname{Ker}\big((u-\lambda_i \mathrm{Id}_E)^{\alpha_i}\big)$ et u_i est la somme de l'homothétie de rapport λ_i et d'un endomorphisme de F_{λ_i} nilpotent d'indice α_i ;

    ▪ le polynôme minimal π_u de u est : $\pi_u = \prod_{i=1}^{p} \pi_{u_i}$;

  – un sous-espace vectoriel F de E est stable par u si et seulement si son intersection avec chacun des F_{λ_i} est stable par u, et dans ce cas :

$$F = \bigoplus_{i=1}^{p} (F \cap F_{\lambda_i}) \ .$$

Thème 5 - Espaces préhilbertiens réels

[S5.1] Produit scalaire

- Une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E est une application φ de $E \times E$ dans $\mathbb{R}$ linéaire par rapport à chacune des deux variables, c'est-à-dire :
  - pour tout $y \in E$, l'application $x \mapsto \varphi(x,y)$ est une forme linéaire sur E ;
  - pour tout $x \in E$, l'application $y \mapsto \varphi(x,y)$ est une forme linéaire sur E.

  La forme bilinéaire $\varphi : E^2 \to \mathbb{R}$ est dite symétrique si :
  $$\forall\,(x,y) \in E^2,\ \varphi(x,y) = \varphi(y,x)\,.$$
  Si φ est une application de E^2 dans $\mathbb{R}$ symétrique et linéaire par rapport à l'une des variables, alors φ est bilinéaire.
- On appelle produit scalaire sur un $\mathbb{R}$-espace vectoriel E une forme bilinéaire symétrique φ qui est, de plus :
  - positive, c'est-à-dire : $\forall\, x \in E,\ \varphi(x,x) \geqslant 0$;
  - et définie, c'est-à-dire : $\forall\, x \in E,\ \varphi(x,x) = 0 \implies x = 0_E$.

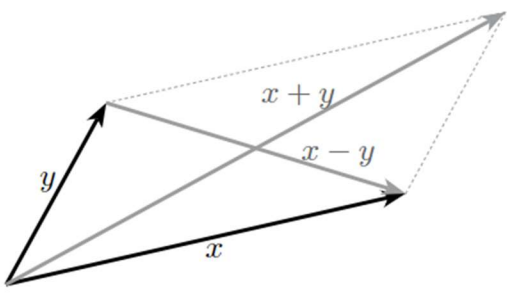
  Dire que φ est définie positive peut aussi s'écrire directement :
  $$\forall\, x \in E \setminus \{0\}\,,\ \varphi(x,x) > 0\,.$$
  Si φ est un produit scalaire sur un $\mathbb{R}$-espace vectoriel E, le couple (E,φ) s'appelle un espace préhilbertien réel.
  Le produit scalaire de deux vecteurs x et y se note en général $\langle x \,|\, y\rangle$.
- **Exemples**
  - Dans $\mathbb{R}^n$, pour $x = (x_1,\dots,x_n)$ et $y = (y_1,\dots,y_n)$, on peut poser :
    $$\langle x \,|\, y\rangle = x_1y_1 + \cdots + x_ny_n\,.$$
    Il s'agit du produit scalaire canonique sur $\mathbb{R}^n$.
    En notant $X \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne ${}^{\mathrm{t}}\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ et $Y \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne ${}^{\mathrm{t}}\begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}$, on a aussi : $\langle x \,|\, y\rangle = {}^{\mathrm{t}}XY$.

    ✓ Dans l'écriture ci-dessus, on assimile la matrice ${}^{\mathrm{t}}XY$, de type (1,1), à un nombre réel.
  - Dans $\mathcal{M}_n(\mathbb{R})$, si $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ et $B = (b_{ij})_{1\leqslant i,j\leqslant n}$, on peut poser :
    $$\langle A \,|\, B\rangle = \mathrm{tr}({}^{\mathrm{t}}AB) = \sum_{1\leqslant i,j\leqslant n} a_{ij}b_{ij}\,.$$
    Il s'agit du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$.

- Dans l'espace vectoriel $\mathscr{C}([a\,;b],\mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$, on peut poser :
$$\langle f\,|\,g\rangle = \int_a^b f(t)g(t)\,\mathrm{d}t\,.$$
On remarquera que ce n'est pas un produit scalaire sur l'espace vectoriel $\mathscr{CM}([a\,;b]\mathbb{R})$ des fonctions continues par morceaux sur $[a\,;b]$.
- Soit I un intervalle quelconque de $\mathbb{R}$. On note $\mathscr{L}^2_c(I,\mathbb{R})$ l'ensemble des fonctions continues et de carré intégrable sur I, c'est-à-dire telles que l'intégrale $\int_I f^2$ converge.
Alors $\mathscr{L}^2_c(I,\mathbb{R})$ est un sous-espace vectoriel de $\mathscr{C}(I,\mathbb{R})$, qui peut être muni d'un produit scalaire par : $\langle f\,|\,g\rangle = \int_I fg$.
- On note $\ell^2(\mathbb{N},\mathbb{R})$ l'ensemble des suites réelles $(u_n)_{n\in\mathbb{N}}$ telles que la série $\sum_{n\in\mathbb{N}} u_n^2$ converge (suites de carré sommable). Alors $\ell^2(\mathbb{N},\mathbb{R})$ est un sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$, qui peut être muni du produit scalaire :
$$\langle u\,|\,v\rangle = \sum_{n=0}^{\infty} u_n v_n\,.$$

[S5.2] Norme associée à un produit scalaire

- Si E est un espace préhilbertien réel, où le produit scalaire est noté $\langle\cdot\,|\,\cdot\rangle$, on appelle norme (euclidienne) associée l'application :
$$N\colon x\in E\mapsto \|x\| = \sqrt{\langle x\,|\,x\rangle}$$
(cela a bien un sens car $\langle x\,|\,x\rangle \geqslant 0$).
- **Propriétés**
  - $\forall\, x\in E\,,\ \|x\| = 0_{\mathbb{R}} \iff x = 0_E\,.$
  - $\forall\, x\in E,\ \forall\,\lambda\in\mathbb{R},\ \|\lambda x\| = |\lambda|\,\|x\|\,.$
  - Un vecteur x est dit unitaire si $\|x\| = 1$.
  Si x est un vecteur non nul, alors le vecteur $\dfrac{x}{\|x\|}$ est unitaire.
  - $\forall\,(x,y)\in E^2,\ \|x+y\|^2 = \|x\|^2 + 2\,\langle x\,|\,y\rangle + \|y\|^2\,.$
  - $\forall\,(x,y)\in E^2,\ \langle x+y\,|\,x-y\rangle = \|x\|^2 - \|y\|^2\,.$
  - *Identité du parallélogramme*
$$\forall\,(x,y)\in E^2,\ \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$



– *Identités de polarisation*

$$\forall\,(x,y)\in E^2,\quad \begin{cases}\langle x\,|\,y\rangle &= \dfrac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right)\\[2ex] \langle x\,|\,y\rangle &= \dfrac{1}{2}\left(\|x+y\|^2-\|x\|^2-\|y\|^2\right).\end{cases}$$

- **Inégalité de Cauchy-Schwarz**
  Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :
  $$|\langle x\,|\,y\rangle| \leqslant \|x\|\,.\,\|y\|$$
  et il y a égalité si et seulement si le système $\{x,y\}$ est lié.
  Remarque : pour f, $g\in\mathscr{L}^2_c(I,\mathbb{R})$ muni du produit scalaire usuel, l'inégalité de Cauchy-Schwarz s'écrit :
  $$\left(\int_I fg\right)^2 \leqslant \left(\int_I f^2\right)\left(\int_I g^2\right).$$
  ✓ L'inégalité de Cauchy-Schwarz implique que le produit scalaire est une application bilinéaire continue sur $E\times E$, lorsque E est muni de la norme associée au produit scalaire (cf. chapitre sur les espaces vectoriels normés).
- **Inégalité de Minkowski (ou inégalité triangulaire)**
  Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :
  $$\|x+y\| \leqslant \|x\| + \|y\|$$
  et il y a égalité si et seulement si la famille $\{x,y\}$ est positivement liée (c'est-à-dire qu'il existe un réel λ *positif* tel que $x=\lambda y$ ou $y=\lambda x$).
  On en déduit :
  $$\forall\,(x,y)\in E^2,\ \big|\,\|x\|-\|y\|\,\big| \leqslant \|x-y\|\,.$$
- Soit E un espace préhilbertien réel. La norme associée au produit scalaire sur E munit E d'une structure d'espace vectoriel normé.

[S5.3] Orthogonalité

Dans toute la suite, E désigne un espace préhilbertien réel.

- **Vecteurs orthogonaux**
  On dit que deux vecteurs x et y de E sont orthogonaux, et on note $x\perp y$, lorsque $\langle x\,|\,y\rangle = 0$.
  La relation $\langle x\,|\,y\rangle = 0$ est équivalente à $\langle y\,|\,x\rangle = 0$. La relation d'orthogonalité est donc symétrique.
- **Orthogonal d'une partie**
  Soit A une partie non vide d'un espace préhilbertien E.
  Un vecteur de E est dit orthogonal à A s'il est orthogonal à tout vecteur de A.
  L'ensemble des vecteurs orthogonaux à A s'appelle l'orthogonal de A, noté $A^\perp$.
  Ainsi :
  $$A^\perp = \{x\in E\mid \forall\, a\in A,\ \langle x\,|\,a\rangle = 0\}\,.$$
- Si x est un vecteur non nul de E, alors $\{x\}^\perp$ est un hyperplan de E.
  Un supplémentaire de $\{x\}^\perp$ est la droite vectorielle $\mathbb{K}.x$.
- Si A est une partie non vide de E, alors $A^\perp$ est un sous-espace vectoriel de E.

- **Propriétés de l'orthogonal**
  - $E^\perp = \{0\}$ et $\{0\}^\perp = E$.
  - $\forall\,(A,B) \in \mathcal{P}(E)^2,\ A \subset B \implies B^\perp \subset A^\perp$.
  - $\forall\,A \in \mathcal{P}(E),\ A^\perp = \mathrm{Vect}(A)^\perp$.
    Il en résulte que, si F est un sous-espace vectoriel de E et si $(e_i)_{i\in I}$ est une base de F, un vecteur x appartient à $F^\perp$ si et seulement si pour tout $i \in I$ on a $\langle x \,|\, e_i\rangle = 0$.
  - $\forall\,A \in \mathcal{P}(E),\ A \subset \left(A^\perp\right)^\perp$.
    ✓ Si F est un sous-espace vectoriel de E, on n'a pas nécessairement $F = \left(F^\perp\right)^\perp$. Un contre-exemple se trouve en **[SF5.7]**.
  - Si F et G sont des sous-espaces vectoriels de E : $(F+G)^\perp = F^\perp \cap G^\perp$.
- Deux sous-espaces vectoriels F et G d'un espace préhilbertien E sont dits orthogonaux, et on note $F \perp G$, lorsque :
$$\forall\,(x,y) \in F \times G, \quad \langle x \,|\, y\rangle = 0\,,$$
autrement dit lorsque tout vecteur de F est orthogonal à tout vecteur de G.
Cela équivaut à : $F \subset G^\perp$ ou $G \subset F^\perp$.
Remarque : dans $\mathbb{R}^3$, on pourra ainsi parler d'une droite et d'un plan orthogonaux, mais deux plans P_1 et P_2 ne peuvent pas être orthogonaux.
Cependant, si les droites $P_1^\perp$ et $P_2^\perp$ sont orthogonales, les plans P_1 et P_2 sont dits perpendiculaires.
- Si F et G sont deux sous-espaces vectoriels orthogonaux de E, alors $F \cap G = \{0\}$.
Plus généralement, des sous-espaces vectoriels orthogonaux deux à deux sont en somme directe.

[S5.4] Familles orthogonales, orthonormales

- Une famille $(x_i)_{i\in I}$ de vecteurs d'un espace préhilbertien réel E est dite orthogonale si : $\forall\,(i,j) \in I^2,\ i \neq j \implies \langle x_i \,|\, x_j\rangle = 0$.
Elle est dite orthonormale si, de plus, $\|x_i\| = 1$ pour tout $i \in I$, c'est-à-dire si :
$$\forall\,(i,j) \in I^2,\ \langle x_i \,|\, x_j\rangle = \delta_{ij}\,.$$
Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs non nuls, alors la famille des vecteurs $\left(\dfrac{x_i}{\|x_i\|}\right)_{i\in I}$ est orthonormale.
- **Relation de Pythagore**
Si $(x_1,\ldots,x_p)$ est une famille orthogonale d'un espace préhilbertien réel, alors :
$$\left\|\sum_{i=1}^{p} x_i\right\|^2 = \sum_{i=1}^{p} \|x_i\|^2\,.$$
Cas particulier : le théorème de Pythagore
Soit E un espace préhilbertien réel, et x, y deux vecteurs de E. Alors :
$$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2\,.$$
- Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs *non nuls*, alors cette famille est libre. En particulier, toute famille orthonormale est libre.

[S5.5] Bases orthonormales

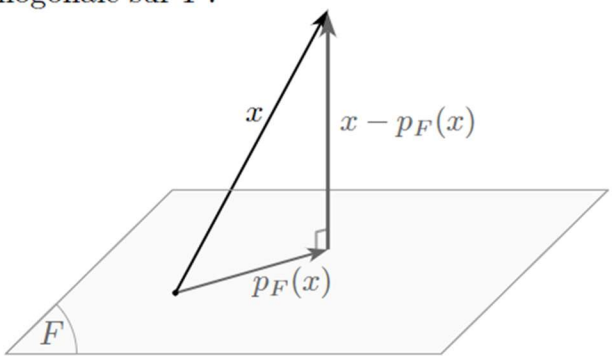
- Un espace vectoriel euclidien est un espace préhilbertien réel de dimension finie.
  Si E est un espace euclidien, on appelle base orthonormale de E toute base de E qui est aussi une famille orthonormale.
  - Pour qu'une famille $(x_1,\dots,x_n)$ de vecteurs d'un espace euclidien forme une base orthonormale, il faut et il suffit que la famille soit orthonormale et que $n = \dim E$.
  - Tout espace euclidien (non réduit à $\{0\}$) possède une base orthonormale.
- **Exemples**
  - Dans $\mathbb{R}^n$ muni du produit scalaire canonique, la base canonique est une base orthonormale.
  - Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, la base canonique $(E_{ij})_{1\leqslant i,j\leqslant n}$ est une base orthonormale.
- **Expression du produit scalaire dans une base orthonormale**
  - Soit E un espace euclidien et $\mathscr{B} = (e_1,\dots,e_n)$ une base orthonormale de E.
    Les coordonnées $(x_1,\dots,x_n)$ d'un vecteur x dans cette base sont données par :
    $$\forall\, i \in [\![1\,;n]\!],\ x_i = \langle e_i \,|\, x\rangle\ .$$
    Ainsi : $\forall\, x \in E,\ x = \sum_{i=1}^{n} \langle e_i \,|\, x\rangle\, e_i\,.$
  - Soient x et y deux vecteurs de E :
    $$x = \sum_{i=1}^{n} x_i e_i \quad \text{et } y = \sum_{i=1}^{n} y_i e_i \quad \text{avec } (x_1,\dots,x_n) \text{ et } (y_1,\dots,y_n) \in \mathbb{K}^n\,.$$
    En notant $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ et $Y = {}^{\mathrm{t}}\begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}$ les matrices colonnes formées des coordonnées dans $\mathscr{B}$ des vecteurs x et y, on a alors (en assimilant comme d'habitude une matrice de type (1,1) à un nombre réel) :
    $$\langle x \,|\, y\rangle = \sum_{i=1}^{n} x_i y_i = {}^{\mathrm{t}}XY = {}^{\mathrm{t}}YX \quad \text{et} \quad \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}\ , \quad \|x\|^2 = {}^{\mathrm{t}}XX\,.$$

    ✓ Il est important de noter que les formules données ci-dessus ne sont valables que si la base choisie est *orthonormale*.

[S5.6] Projection orthogonale sur un sous-espace de dimension finie

- Soit F un sous-espace vectoriel *de dimension finie* d'un espace préhilbertien E.
  Alors :
  $$E = F \oplus F^{\perp} \qquad \text{et} \qquad \left(F^{\perp}\right)^{\perp} = F\,.$$
  On en déduit que si F et G sont deux sous-espaces vectoriels d'un espace préhilbertien de dimension finie E, alors :
  $$(F \cap G)^{\perp} = F^{\perp} + G^{\perp}\,.$$

- $F^\perp$ s'appelle le supplémentaire orthogonal de F.
  On peut alors considérer la projection p_F sur F parallèlement à $F^\perp$; on l'appelle projection orthogonale sur F.



- **Expression analytique**
  Si $(e_1,\dots,e_p)$ est une base *orthonormale* de F, on a la formule :

$$\forall\, x \in E,\ p_F(x) = \sum_{i=1}^{p} \langle e_i \,|\, x\rangle\, e_i\,.$$

  Si $(e_1,\dots,e_p)$ est seulement une base *orthogonale* de F, on a la formule :

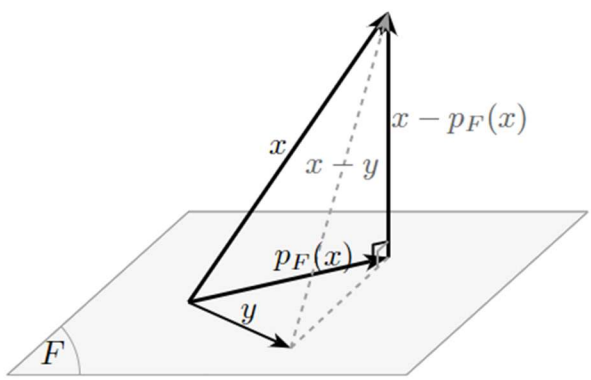
$$\forall\, x \in E,\ p_F(x) = \sum_{i=1}^{p} \frac{\langle e_i \,|\, x\rangle}{\|e_i\|^2}\, e_i\,.$$

  En particulier, le projeté orthogonal sur la droite D de base $a \neq 0$ est :

$$p_D(x) = \frac{\langle a \,|\, x\rangle}{\|a\|^2}\, a\,.$$

- **Lien avec la distance**
  Pour tout x appartenant à E, le vecteur $p_F(x)$ est l'unique vecteur de F réalisant la distance de x à F, c'est-à-dire :

$$\|x - p_F(x)\| = \inf_{y\in F} \|x - y\| = \mathrm{d}(x,F)\,.$$



  D'après le théorème de Pythagore, on a :

$$d(x,F)^2 = \|x\|^2 - \|p_F(x)\|^2\,.$$

En particulier, si l'on connaît une base *orthonormale* $(e_1,\dots,e_p)$ de F, on aura la formule, pour tout $x \in E$:

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{p} \langle e_i \,|\, x\rangle^2.$$

Si la base $(e_1,\dots,e_p)$ est seulement *orthogonale*, on aura :

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{p} \frac{\langle e_i \,|\, x\rangle^2}{\|e_i\|^2}\,.$$

Des formules précédentes on déduit immédiatement **l'inégalité de Bessel**.
Si $(e_1,\dots,e_n)$ est une famille orthonormale d'un espace préhilbertien E, pour tout $x \in E$ on a :

$$\sum_{i=1}^{n} \langle e_i \,|\, x\rangle^2 \leqslant \|x\|^2\,.$$

[S5.7] Procédé d'orthonormalisation de Gram-Schmidt

- Soit E un espace préhilbertien réel.
  On se donne dans E une famille *libre* de vecteurs $(a_1,a_2,\dots,a_n,\dots)$, finie ou non.
  Il existe alors une et une seule famille *orthonormale* $(e_1,e_2,\dots,e_n,\dots)$ telle que :
  (i) pour tout $k \geqslant 1$, $\mathrm{Vect}(e_1,\dots,e_k) = \mathrm{Vect}(a_1,\dots,a_k)$;
  (ii) pour tout $k \geqslant 1$, $\langle e_k \,|\, a_k\rangle > 0$.
  Dans ce théorème, la condition (ii) ne sert qu'à assurer l'unicité des e_k.
- Le procédé de construction est tout aussi important à retenir que le théorème. Les vecteurs e_k se calculent par récurrence de la façon suivante :
  – on pose $e_1 = \dfrac{a_1}{\|a_1\|}$;
  – si l'on a calculé $e_1,\dots,e_k$ alors on considère le vecteur :

$$b_{k+1} = a_{k+1} - \sum_{i=1}^{k} \langle e_i \,|\, a_{k+1}\rangle\, e_i\,,$$

  puis l'on pose $e_{k+1} = \dfrac{b_{k+1}}{\|b_{k+1}\|}$.

Pour retenir facilement ces formules, il suffit de remarquer que le vecteur $\sum_{i=1}^{k} \langle e_i \,|\, a_{k+1}\rangle\, e_i$ n'est autre que le projeté orthogonal $p(a_{k+1})$ de a_{k+1} sur le sous-espace vectoriel $\mathrm{Vect}(e_1,\dots,e_k) = \mathrm{Vect}(a_1,\dots,a_k)$, d'après la formule qui a été rappelée précédemment. Le vecteur $b_{k+1} = a_{k+1} - p(a_{k+1})$ appartient donc bien au supplémentaire orthogonal de $\mathrm{Vect}(a_1,\dots,a_k)$ dans $\mathrm{Vect}(a_1,\dots,a_k,a_{k+1})$.
Dans de nombreux cas, trouver une base orthogonale (et non orthonormale) suffit. On trouve alors une telle base $(e'_1,\dots,e'_n)$ comme suit :

– on pose $e'_1 = a_1$;

– si l'on a calculé $e'_1, \dots, e'_k$ alors on pose :

$$e'_{k+1} = a_{k+1} - \sum_{i=1}^{k} \frac{\langle e'_i \,|\, a_{k+1} \rangle}{\|e'_i\|^2} e'_i \,.$$

- **Interprétation matricielle**
  Supposons maintenant E de dimension finie n, et soit $\mathscr{B} = (a_1, \dots, a_n)$ une base de E. Le procédé ci-dessus a permis de construire une base orthonormale $\mathscr{B}' = (e_1, \dots, e_n)$.
  On peut alors vérifier que la matrice de passage de la base $\mathscr{B}$ à son orthonormalisée par le procédé de Schmidt $\mathscr{B}'$ est une matrice triangulaire supérieure à éléments diagonaux strictement positifs.

[S5.8] Suites totales

- Soit E un espace préhilbertien réel. Une suite $\mathscr{B} = (e_n)_{n\in\mathbb{N}}$ d'éléments de E est dite totale si $E = \overline{\mathrm{Vect}(\mathscr{B})}$, autrement dit si $\mathrm{Vect}(\mathscr{B})$ est dense dans l'espace vectoriel normé $(E, \|\cdot\|)$.
- Si $\mathscr{B}$ est une suite totale alors $\big(\mathrm{Vect}(\mathscr{B})\big)^{\perp} = \{0\}$.

  ✓ La réciproque de cette propriété est fausse : si $\big(\mathrm{Vect}(\mathscr{B})\big)^{\perp} = \{0\}$, la famille $\mathscr{B}$ n'est pas forcément totale.

  ✓ Cette propriété implique que, lorsque E est de dimension finie, une famille totale $\mathscr{B}$ est nécessairement génératrice de E.
- Soit $\mathscr{B} = (e_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace préhilbertien réel E. Pour tout $n \in \mathbb{N}$, notons p_n la projection orthogonale sur le sous-espace vectoriel $\mathrm{Vect}(e_0, \dots, e_n)$.
  La suite $\mathscr{B}$ est totale si et seulement si pour tout $x \in E$ la suite $\big(p_n(x)\big)_{n\in\mathbb{N}}$ converge vers x (dans l'espace vectoriel normé $(E, \|\cdot\|)$).
- Soit $\mathscr{B} = (e_n)_{n\in\mathbb{N}}$ une suite *orthonormale* d'éléments d'un espace préhilbertien E.
  – Pour tout $x \in E$ la série $\sum_{n\in\mathbb{N}} \langle e_n \,|\, x\rangle^2$ converge et :

  $$\sum_{n=1}^{+\infty} \langle e_n \,|\, x\rangle^2 \leqslant \|x\|^2 \quad \text{(inégalité de Bessel généralisée).}$$

  – La suite $\mathscr{B}$ est totale si et seulement si pour tout $x \in E$, $\|x\|^2 = \sum_{n=1}^{+\infty} \langle e_n \,|\, x\rangle^2$ (égalité de Parseval).
- **Polynômes orthogonaux**
  Soit $E = \mathscr{C}([a\,;b],\mathbb{R})$ l'espace vectoriel des fonctions continues sur $[a\,;b]$ $(a < b)$, et soit ω une fonction continue positive sur $[a\,;b]$ ne s'annulant qu'en un nombre fini de points (ω s'appelle une fonction de poids).

– L'application $\varphi\colon (f,g) \mapsto \displaystyle\int_a^b f(t)g(t)\omega(t)\,\mathrm{d}t$ est un produit scalaire sur E.

– Il existe une et une seule suite $(P_n)_{n\in\mathbb{N}}$ de polynômes unitaires, orthogonaux deux à deux pour ce produit scalaire, et tels que, pour tout $n \in \mathbb{N}$, $\deg(P_n) = n$.

– La suite $(P_n)_{n\in\mathbb{N}}$ est une suite totale de l'espace préhilbertien (E,φ).

[S5.9] Caractérisation des projecteurs orthogonaux et des symétries orthogonales

E désigne toujours un espace préhilbertien réel.

- **Caractérisation des projecteurs orthogonaux parmi les projecteurs**
  Soit p un *projecteur* de E ($p \in \mathscr{L}(E)$ et $p\circ p = p$). Les trois propriétés suivantes sont équivalentes :

  (i) p est un projecteur orthogonal (c'est-à-dire $\operatorname{Im} p \perp \operatorname{Ker} p$),

  (ii) pour tout $x \in E$, $\|p(x)\| \leqslant \|x\|$,

  (iii) pour tous x, $y \in E$, $\langle p(x)\,|\,y\rangle = \langle x\,|\,p(y)\rangle$.

  ✓ La propriété **(ii)** implique que p est une application linéaire continue sur sur E, lorsque E est muni de la norme associée au produit scalaire (cf. chapitre sur les espaces vectoriels normés).

- **Caractérisation des symétries orthogonales parmi les symétries**
  Soit p une *symétrie* de E ($s \in \mathscr{L}(E)$ et $s\circ s = \mathrm{Id}_E$). Les trois propriétés suivantes sont équivalentes :

  (i) s est une symétrie orthogonale,

  (ii) pour tout $x \in E$, $\|s(x)\| = \|x\|$,

  (iii) pour tous x, $y \in E$, $\langle s(x)\,|\,y\rangle = \langle x\,|\,s(y)\rangle$.

Thème 6 - Endomorphismes des espaces euclidiens

Dans tout ce chapitre, E désigne un espace euclidien, c'est-à-dire un espace préhilbertien réel de dimension finie, où le produit scalaire est noté $\langle \cdot | \cdot \rangle$.

[S6.1] Isométries vectorielles

- Une isométrie est un endomorphisme u de E tel que :

$$\forall\,(x,y) \in E^2,\ \langle u(x) \,|\, u(y)\rangle = \langle x \,|\, y\rangle\ .$$

  On dit que u conserve le produit scalaire.

- Une isométrie u peut aussi être caractérisée par l'une des propriétés équivalentes suivantes :

  (i) $\forall\, x \in E,\ \|u(x)\| = \|x\|$ (on dit que u conserve la norme),

  (ii) u transforme une/toute base orthonormale en une base orthonormale.

  La propriété (ii) implique que, si u est une isométrie, u est aussi un endomorphisme bijectif (pour cette raison, une isométrie est aussi appelée automorphisme orthogonal).

- **Exemple :** une symétrie orthogonale par rapport à un sous-espace vectoriel F de E est une isométrie.

- **Matrices orthogonales**

  . Une matrice carrée $A \in \mathcal{M}_n(\mathbb{R})$ est dite orthogonale si c'est la matrice d'une isométrie *dans une base orthonormale*, ou encore si c'est la matrice de passage d'une base orthonormale à une base orthonormale.

  Cela équivaut à dire que ${}^tAA = A{}^tA = I_n$.

  Pour reconnaître rapidement si une matrice est orthogonale, il est important de se souvenir de la propriété suivante :

  > une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ses vecteurs colonnes (ou ses vecteurs lignes), considérés comme éléments de $\mathbb{R}^n$, forment une famille orthonormale de $\mathbb{R}^n$ pour le produit scalaire canonique.

  . L'ensemble $\mathcal{O}_n(\mathbb{R})$ des matrices carrées orthogonales d'ordre n est un sous-groupe du groupe $(\mathrm{GL}_n(\mathbb{R}),\times)$ des matrices inversibles.

  . Si A est une matrice orthogonale, toute valeur propre λ de A, réelle ou complexe, est telle que $|\lambda| = 1$.

  En particulier, les seules valeurs propres possibles d'une isométrie sont ± 1.

- **Groupe orthogonal**

  . L'ensemble $\mathcal{O}(E)$ des isométries de E est un groupe pour la loi $\circ$; c'est un sous-groupe de $(\mathrm{GL}(E),\circ)$, appelé groupe orthogonal.

- Si $u \in \mathcal{O}(E)$, alors $|\det u| = 1$.

  L'ensemble des isométries de déterminant égal à $+1$ est un sous-groupe de $\mathcal{O}(E)$, appelé groupe spécial orthogonal et noté $\mathcal{O}^+(E)$; ses éléments sont appelés des rotations ou isométries positives.

  Une symétrie orthogonale par rapport à un hyperplan de H s'appelle une réflexion ; c'est une isométrie de déterminant -1.

  Une isométrie de déterminant -1 s'appelle aussi une isométrie négative ; leur ensemble est noté $\mathcal{O}^-(E)$.

[S6.2] Produit vectoriel

- **Produit mixte**

  Soit E un espace vectoriel euclidien orienté de dimension n, et $(x_1,\dots,x_n)$ une famille de vecteurs de E.

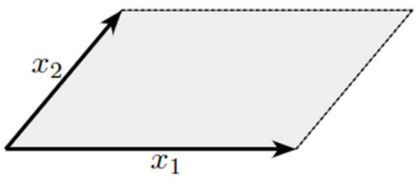
  Alors le déterminant de cette famille dans *toute* base orthonormale directe est le même ; il s'appelle le produit mixte de $(x_1,\dots,x_n)$ et est noté $[x_1,\dots,x_n]$:

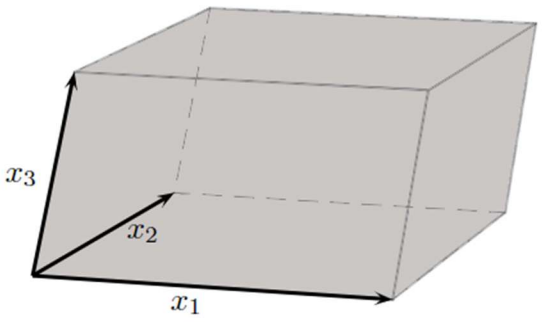
  $$[x_1,\dots,x_n] = \det_{\mathscr{B}}(x_1,\dots,x_n) \quad \text{où } \mathscr{B} \text{ est une base orthonormale directe.}$$

  Interprétation géométrique

  En dimension 2, la valeur absolue du produit mixte de x_1 et x_2 représente l'aire du parallélogramme construit sur x_1 et x_2.

  En dimension 3, la valeur absolue du produit mixte de trois vecteurs x_1, x_2 et x_3 représente le volume du parallélépipède construit sur ces trois vecteurs.





- **Produit vectoriel**

  Soit E_3 un espace vectoriel euclidien orienté de dimension 3, et u, v deux vecteurs de E_3. Alors il existe un et un seul vecteur $a \in E_3$ tel que :

  $$\forall\, w \in E_3,\ [u,v,w] = \langle a \,|\, w \rangle\,.$$

  a s'appelle le produit vectoriel de u et v et est noté $u \wedge v$. Il est donc caractérisé par :

  $$\forall\, (u,v,w) \in E_3^3,\ [u,v,w] = \langle u \wedge v \,|\, w \rangle\,.$$

- **Propriétés**

  - Soient u et v deux vecteurs de E_3.
    - $u \wedge v = 0_E \iff$ la famille (u,v) est liée.

– Le vecteur $u \wedge v$ appartient à $\mathrm{Vect}(u,v)^\perp$. De plus, si la famille (u,v) est libre, la famille $(u,v,u \wedge v)$ est une base directe.

- L'application $\begin{cases} E_3 \times E_3 & \longrightarrow & E_3 \\ (u,v) & \longmapsto & u \wedge v \end{cases}$ est bilinéaire antisymétrique.
- Si u et v ont pour coordonnées respectives (x,y,z) et (x',y',z') dans une base orthonormale directe $\mathscr{B}$, les coordonnées dans cette même base du produit vectoriel $u \wedge v$ sont :

$$\begin{vmatrix} y & y' \\ z & z' \end{vmatrix}, \quad -\begin{vmatrix} x & x' \\ z & z' \end{vmatrix} \quad \text{et} \quad \begin{vmatrix} x & x' \\ y & y' \end{vmatrix}.$$

- *Formule du double produit vectoriel*

$$\forall\,(u,v,w) \in E_3^3,\ u \wedge (v \wedge w) = \langle u \,|\, w \rangle\, v - \langle u \,|\, v \rangle\, w\,.$$

[S6.3] Classification des isométries en dimension 2

On note ici E un plan vectoriel euclidien, muni d'une base orthonormale $\mathscr{B}$, et on supposera E orienté par le choix de cette base.
Soit u une isométrie de E. Il y a deux cas possibles.

- $\det u = +1$

  u est alors une rotation. Il existe un réel $\theta \in\,]-\pi\,;\pi]$, appelé angle de la rotation, tel que la matrice de u dans *toute* base orthonormale *directe* soit de la forme :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

  Sa matrice dans une base orthonormale *indirecte* s'obtient en changeant θ en $-\theta$.

- $\det u = -1$

  Il existe alors un réel θ tel que :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

  mais ici, θ dépend de la base orthonormale $\mathscr{B}$ choisie.

  Il est alors facile de vérifier, en cherchant les vecteurs propres de u pour les valeurs propres ± 1, que u est une symétrie orthogonale par rapport à une droite D (réflexion). Cette droite est engendrée par le vecteur de coordonnées $\left(\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\right)$ dans la base $\mathscr{B}$.

[S6.4] Classification des isométries en dimension 3

On note ici E un espace vectoriel euclidien orienté de dimension 3.
Soit u une isométrie de E. Il y a deux cas possibles.

- $\det u = +1$

  u est une rotation. On cherche alors l'ensemble des vecteurs invariants par u, c'est-à-dire le sous-espace propre $E_1(u)$ associé à la valeur propre 1. On montre qu'il n'y a que deux cas possibles :

  – soit $\dim E_1(u) = 3$: u est alors l'application identique de E.

  – soit $\dim E_1(u) = 1$: l'ensemble des vecteurs invariants de u est ici une droite D, appelée *axe de la rotation.*

  Le plan $P = D^\perp$ est alors stable par u ; en considérant un vecteur de base unitaire e_1 de l'axe D et une base orthonormale (e_2,e_3) de P telle que la base $\mathscr{B} = (e_1,e_2,e_3)$ soit directe, la matrice de u dans $\mathscr{B}$ est de la forme :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

  θ est l'angle de la rotation ; il vérifie la relation : $\operatorname{tr} u = 1 + 2\cos\theta$ (on rappelle que la trace d'un endomorphisme ne dépend pas de la base dans laquelle on écrit sa matrice). Cette relation très simple permet de déterminer facilement $\pm\theta$. Pour déterminer la valeur exacte de θ, dont le signe dépend de l'orientation, nous vous invitons à vous reporter à l'exemple traité en **[SF6.10]** page 364.

- $\det u = -1$

  On montre alors qu'il existe une base orthonormale $\mathscr{B} = (e_1,e_2,e_3)$ (que l'on peut choisir directe) telle que la matrice de u dans cette base soit de la forme :

$$M_{\mathscr{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

  – si $\theta \equiv 0 \pmod{2\pi}$, u est la symétrie orthogonale par rapport au plan P de base (e_2,e_3). Ce plan est le sous-espace propre $E_1(u)$.

  – si $\theta \equiv \pi \pmod{2\pi}$, on a alors simplement $u = -\mathrm{Id}_E$.

  – sinon, comme on peut le vérifier facilement à l'aide de la représentation matricielle, u est la composée commutative de la réflexion par rapport à P et de la rotation d'axe $\mathbb{R}e_1$ et d'angle θ ; cet axe est alors l'ensemble des vecteurs x tels que $u(x) = -x$, c'est-à-dire le sous-espace propre $E_{-1}(u)$, et l'ensemble des vecteurs invariants de u est réduit à $\{0\}$. L'angle θ de la rotation vérifie ici la relation : $\operatorname{tr} u = -1 + 2\cos\theta$.

  u s'appelle l'antirotation d'axe $\mathbb{R}e_1$ et d'angle θ.

  ✓ Vous remarquerez que la simple détermination des sous-espaces propres de u et le calcul de sa trace permettent de caractériser u presque entièrement (au signe près de θ dans le cas d'une rotation ou d'une antirotation).

Les isométries sont en fait caractérisées par la dimension du sous-espace des vecteurs invariants, comme le montrent les tableaux récapitulatifs suivants.

- si u est une isométrie en dimension 2

$\dim E_1(u)$	Nature de u	Type d'isométrie
0	rotation	$+$
1	réflexion d'axe $E_1(u)$	$-$
2	Id_E	$+$

- si u est une isométrie en dimension 3

$\dim E_1(u)$	Nature de u	Type d'isométrie
0	antirotation d'axe $E_{-1}(u)$	$-$
1	rotation d'axe $E_1(u)$	$+$
2	réflexion par rapport au plan $E_1(u)$	$-$
3	Id_E	$+$

[S6.5] Réduction d'une isométrie dans le cas général

On note ici E un espace vectoriel euclidien de dimension $n \in \mathbb{N}^*$.
On rappelle que si $u \in \mathcal{O}(E)$, alors $\det u = \pm 1$ et $\mathrm{Sp}(u) \subset \{-1,1\}$.

- **Sous-espaces stables**
  Soit u une isométrie de E. Si F est un sous-espace de E stable par u, alors $F^\perp$ est aussi stable par u.

- **Réduction d'une isométrie dans une base orthonormale**
  Soit u une isométrie de E.
  Il existe une base orthonormale $\mathscr{B}$ de E telle que la matrice A de u dans $\mathscr{B}$ soit diagonale par blocs, de la forme :

$$A = \begin{bmatrix} \mathrm{I}_p & 0 & \dots & \dots & 0 \\ 0 & -\mathrm{I}_q & & & \vdots \\ \vdots & & R_1 & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \dots & \dots & 0 & R_s \end{bmatrix}$$

  où I_p et I_q désignent les matrices identités d'ordres p et q et où pour tout $i \in [\![1\,;s]\!]$ R_i est une matrice de rotation de la forme $R_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$ avec $\theta_i \in \mathbb{R} \setminus \pi\mathbb{Z}$.

  ✓ Dans l'écriture ci-dessus, certains blocs peuvent ne pas être présents. Par exemple, si les blocs R_i ne figurent pas, u est une symétrie orthogonale, et il s'agit du seul cas où une isométrie peut être diagonalisable.

[S6.6] Endomorphismes symétriques

- Un endomorphisme u d'un espace vectoriel euclidien E est dit symétrique (relativement au produit scalaire considéré) si :

$$\forall\,(x,y)\in E^2,\ \langle u(x)\,|\,y\rangle = \langle x\,|\,u(y)\rangle\,.$$

✓ Cette notion dépend du produit scalaire considéré sur E.

L'ensemble $\mathscr{S}(E)$ des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathscr{L}(E)$.

- Un endomorphisme est symétrique si et seulement si sa matrice dans une/toute base *orthonormale* est symétrique.
- Un projecteur est un endomorphisme symétrique si et seulement si c'est un projecteur orthogonal.
- Une symétrie est un endomorphisme symétrique si et seulement si c'est une symétrie orthogonale.
- Soit u un endomorphisme symétrique de E.
  Si F est un sous-espace vectoriel de E stable par u, alors $F^\perp$ est aussi stable par u.
- **Théorème spectral**
  Soit E un espace vectoriel euclidien, et u un endomorphisme symétrique de E.
  (i) Le polynôme caractéristique de u est scindé dans $\mathbb{R}[X]$.
  (ii) Les sous-espaces propres de u sont orthogonaux deux à deux.
  (iii) Il existe une base orthonormale de E formée de vecteurs propres de u (ce que l'on traduit en disant brièvement : l'endomorphisme u est diagonalisable dans une base orthonormale).

  Puisque la matrice d'un endomorphisme symétrique dans une base orthonormale est une matrice symétrique, on en déduit le théorème suivant.
  Si S est une matrice symétrique réelle, il existe une matrice diagonale D et une matrice orthogonale P telles que :

$$S \;=\; PDP^{-1} \;=\; PD^tP$$

(on dit aussi que S est orthodiagonalisable).

Thème 7 - Espaces vectoriels normés

Dans tout ce chapitre, les espaces vectoriels considérés sont des espaces vectoriels sur $\mathbb{K}$, où $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Si $\lambda \in \mathbb{K}$, l'écriture $|\lambda|$ désigne la *valeur absolue* de λ si $\mathbb{K} = \mathbb{R}$, et le *module* de λ si $\mathbb{K} = \mathbb{C}$.

[S7.1] Espaces vectoriels normés

- **Normes sur un espace vectoriel**

  Soit E un $\mathbb{K}$-espace vectoriel. On appelle norme sur E une application $N : E \longrightarrow \mathbb{R}$ telle que :

  (i) $\forall\, x \in E,\ N(x) \geqslant 0$ (*positivité*) ;

  (ii) $\forall\, x \in E,\ N(x) = 0 \implies x = 0$ (*séparation*) ;

  (iii) $\forall\, x \in E,\ \forall\, \lambda \in \mathbb{K},\ N(\lambda x) = |\lambda| N(x)$ (*homogénéité*) ;

  (iv) $\forall\, (x,y) \in E^2,\ N(x+y) \leqslant N(x) + N(y)$ (*inégalité triangulaire*).

  On dit alors que le couple (E,N) est un espace vectoriel normé. On écrit souvent $N(x) = \|x\|$.

  Un vecteur x d'un espace vectoriel normé $(E, \|\cdot\|)$ est dit unitaire si $\|x\| = 1$.

  Pour tout vecteur $x \in E$ non nul, $\dfrac{x}{\|x\|}$ est unitaire.

- **Inégalité triangulaire**

  Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Alors :

  – $\forall\, (x,y) \in E^2,\ \big|\, \|x\| - \|y\| \,\big| \leqslant \|x+y\| \leqslant \|x\| + \|y\|$.

  – $\forall\, (x_1,\ldots,x_n) \in E^n,\ \forall\, (\lambda_1,\ldots,\lambda_n) \in \mathbb{K}^n,\ \left\| \sum_{i=1}^{n} \lambda_i x_i \right\| \leqslant \sum_{i=1}^{n} |\lambda_i|\, \|x_i\|$.

- **Exemples**

  . La valeur absolue dans $\mathbb{R}$ et le module dans $\mathbb{C}$ sont des normes.

  . Soit $(E, \langle\cdot\,|\,\cdot\rangle)$ un espace vectoriel préhilbertien réel. Alors la norme euclidienne associée au produit scalaire $\langle\cdot\,|\,\cdot\rangle$, définie par $\|x\| = \sqrt{\langle x\,|\,x\rangle}$, est une norme sur E.

  . Soit E un espace vectoriel de dimension finie n, muni d'une base $\mathscr{B}$.
  Pour tout vecteur x de coordonnées $(x_1,\ldots,x_n) \in \mathbb{K}^n$ dans $\mathscr{B}$, on note :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \ , \quad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad \text{et} \quad \|x\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i| \ .$$

  Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont trois normes sur E.

. Soient $a < b$ deux réels. Pour toute fonction $f \in \mathscr{C}([a\,;b],\mathbb{K})$, on note :

$$\|f\|_1 = \int_a^b |f(t)|\,\mathrm{d}t\,, \quad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2\,\mathrm{d}t} \quad \text{et} \quad \|f\|_\infty = \sup_{t\in[a;b]} |f(t)|\,.$$

Alors $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont trois normes sur $\mathscr{C}([a\,;b],\mathbb{K})$, appelées respectivement : norme de la convergence en moyenne, norme de la convergence en moyenne quadratique et norme de la convergence uniforme.

- **Produit d'espaces vectoriels normés**

Soient $(E_1,N_1),\ldots,(E_p,N_p)$ une famille finie de $\mathbb{K}$-espaces vectoriels normés. On peut définir sur l'espace vectoriel produit $E = E_1 \times \cdots \times E_p$ l'application :

$$N : (x_1,\ldots,x_p) \mapsto \max_{1\leqslant i\leqslant p} N_i(x_i)\,.$$

Alors N est une norme sur E ; l'espace vectoriel normé (E,N) s'appelle l'espace vectoriel normé produit des (E_i,N_i).

- **Algèbre normée**

Une algèbre normée est une $\mathbb{K}$-algèbre $(\mathcal{A},+,.,\times)$ munie d'une norme $\|\cdot\|$ telle que :

$$\forall\,(x,y) \in \mathcal{A}^2,\ \|x\times y\| \leqslant \|x\|\,\|y\|\,.$$

On dira qu'il s'agit d'une algèbre normée unitaire si, de plus : $\|1_{\mathcal{A}}\| = 1$.

- **Distance associée à une norme**

Soit $(E,\|\cdot\|)$ un espace vectoriel normé. Pour tout $(x,y) \in E^2$, on appelle distance de x à y le réel $d(x,y) = \|x-y\|$.

Propriétés

– $\forall\,(x,y) \in E^2,\ d(x,y) \geqslant 0$.

– $\forall\,(x,y) \in E^2,\ d(x,y) = d(y,x)$.

– $\forall\,(x,y) \in E^2,\ d(x,y) = 0 \iff x = y$.

– $\forall\,(x,y,z) \in E^3,\ \big|d(x,y) - d(y,z)\big| \leqslant d(x,z) \leqslant d(x,y) + d(y,z)$.

- **Distance à une partie**

Soit $(E,\|\cdot\|)$ un espace vectoriel normé, A une partie de E non vide, et $x \in E$. L'ensemble $\{\|x-a\|\,,\ a \in A\}$ est une partie non vide de $\mathbb{R}$, minorée par 0 ; elle admet donc une borne inférieure, appelée distance de x à A, et notée $d(x,A)$:

$$d(x,A) = \inf_{a\in A} d(x,a)\,.$$

[S7.2] Boules

- Soit $(E,\|\cdot\|)$ un espace vectoriel normé, soit a un vecteur de E et soit r un réel positif. On appelle boule *ouverte* de centre a et de rayon r l'ensemble :

$$\mathcal{B}(a,r) = \{x \in E \mid d(a,x) < r\} = \{x \in E \mid \|a-x\| < r\}\,.$$

On appelle boule *fermée* de centre a et de rayon r l'ensemble :

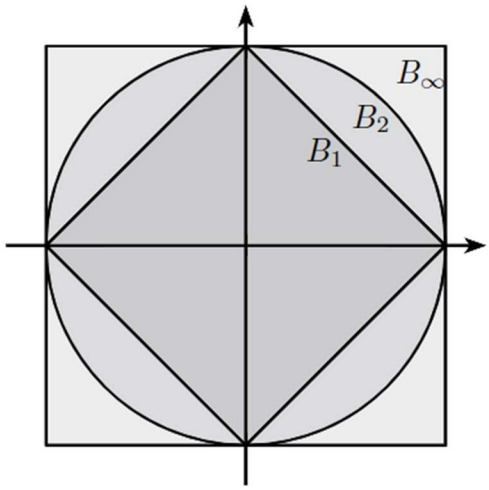
$$\mathcal{B}_f(a,r) = \{x \in E \mid d(a,x) \leqslant r\} = \{x \in E \mid \|a-x\| \leqslant r\}\,.$$

On appelle sphère de centre a et de rayon r l'ensemble :

$$\mathcal{S}(a,r) = \{x \in E \mid d(a,x) = r\} = \{x \in E \mid \|a - x\| = r\}\,.$$

Dans le cas où $a = 0_E$ et $r = 1$, on parle de boule unité et de sphère unité.

- **Exemples**
  - Dans $\mathbb{R}$ muni de la valeur absolue $|\cdot|$:
    $\mathcal{B}(a,r) =]a-r\,;a+r[$, $\mathcal{B}_f(a,r) = [a-r\,;a+r]$ et $\mathcal{S}(a,r) = \{a-r\,,a+r\}$.
  - La figure ci-après représente les boules unités fermées pour les trois normes usuelles de $\mathbb{R}^2$.
    - Pour la norme $\|\cdot\|_1$, $B_1 = \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; |x| + |y| \leqslant 1\right\}$.
    - Pour la norme $\|\cdot\|_2$, $B_2 = \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 \leqslant 1\right\}$.
    - Pour la norme $\|\cdot\|_\infty$, $B_\infty = \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; |x| \leqslant 1 \text{ et } |y| \leqslant 1\right\}$.



- Soit $(E, \|\cdot\|)$ un espace vectoriel normé, soit $(a_1,a_2) \in E^2$, et soient r_1 et r_2 deux réels strictement positifs. On note $d = d(a_1,a_2) = \|a_1 - a_2\|$. Alors :
  - $\mathcal{B}(a_1,r_1) \subset \mathcal{B}(a_2,r_2) \iff d \leqslant r_2 - r_1$.
  - $\mathcal{B}_f(a_1,r_1) \subset \mathcal{B}_f(a_2,r_2) \iff d \leqslant r_2 - r_1$.
  - $\mathcal{B}(a_1,r_1) \cap \mathcal{B}(a_2,r_2) = \emptyset \iff d \geqslant r_1 + r_2$.
  - $\mathcal{B}_f(a_1,r_1) \cap \mathcal{B}_f(a_2,r_2) = \emptyset \iff d > r_1 + r_2$.
- **Parties convexes**

  Soit E un $\mathbb{K}$-espace vectoriel. Pour tout $(x,y) \in E^2$, on appelle segment $[x\,;y]$ l'ensemble :

  $$[x\,;y] = \{tx + (1-t)y \mid t \in [0\,;1]\}\,.$$

  Une partie A d'un espace vectoriel est dite convexe si, pour tout $(x,y) \in A^2$, le segment $[x\,;y]$ est inclus dans A.

  Dans un espace vectoriel normé $(E, \|\cdot\|)$, toute boule (ouverte ou fermée) est une partie convexe de E.
- **Parties bornées**

  Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Une partie A de E est dite bornée si :

  $$\exists M \in \mathbb{R},\ \forall x \in A,\ \|x\| \leqslant M\,.$$

  Autrement dit, A est bornée s'il existe un réel positif M tel que $A \subset \mathcal{B}_f(0,M)$.

[S7.3] Comparaison des normes

- On dit que deux normes N_1 et N_2 sur E sont équivalentes s'il existe deux réels strictement positifs α et β tels que :

$$\forall\, x \in E,\ \alpha N_1(x) \leqslant N_2(x) \leqslant \beta N_1(x)\,,$$

ce qui équivaut à dire que les rapports $\dfrac{N_1}{N_2}$ et $\dfrac{N_2}{N_1}$ sont bornés sur $E \setminus \{0\}$.

- **Exemple**

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, muni d'une base $\mathscr{B}$. On note $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ les normes usuelles associées. Alors :

$$\forall\, x \in E,\ \|x\|_\infty \leqslant \|x\|_1 \leqslant n \cdot \|x\|_\infty \qquad \text{et} \qquad \|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n} \cdot \|x\|_\infty\,.$$

Cela démontre que les trois normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont équivalentes.

- **Caractérisations**

Soit E un $\mathbb{K}$-espace vectoriel, et N_1, N_2 deux normes sur E.

- N_1 et N_2 sont équivalentes si et seulement si toute boule ouverte pour l'une de ces deux normes contient une boule ouverte pour l'autre, de même centre.
- Les deux normes N_1 et N_2 sont équivalentes si et seulement si toute partie bornée pour l'une de ces deux normes est aussi bornée pour l'autre.

- **Équivalence des normes en dimension finie**

Si E est un $\mathbb{K}$-espace vectoriel de *dimension finie*, toutes les normes sur E sont équivalentes.

> ✓ Ce résultat ne subsiste pas en dimension infinie.
> Par exemple, dans $E = \mathscr{C}([a\,;b],\mathbb{R})$, les normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont deux à deux non équivalentes.

[S7.4] Suites d'un espace vectoriel normé

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

- On note $E^{\mathbb{N}}$ l'ensemble des suites définies sur $\mathbb{N}$ à valeurs dans E, c'est-à-dire l'ensemble des applications de la forme $u : \begin{cases} \mathbb{N} & \longrightarrow & E \\ n & \longmapsto & u_n\,. \end{cases}$

Pour toutes suites u, $v \in E^{\mathbb{N}}$ et pour tout scalaire $\lambda \in \mathbb{K}$, on note :

$$u + v : n \longmapsto u_n + v_n \qquad \text{et} \qquad \lambda.u : n \longmapsto \lambda u_n\,.$$

Muni de ces lois, l'ensemble $(E^{\mathbb{N}}, +\,, .)$ est un $\mathbb{K}$-espace vectoriel.

- **Suites bornées**

Soit $u \in E^{\mathbb{N}}$ une suite. On dit que u est bornée si :

$$\exists\, M \in \mathbb{R},\ \forall\, n \in \mathbb{N},\ \|u_n\| \leqslant M\,.$$

Ainsi, u est bornée si l'ensemble $\{u_n \mid n \in \mathbb{N}\}$ est une partie bornée de E.

L'ensemble noté $\ell^\infty(E)$ des suites bornées d'éléments de E est un sous-espace vectoriel de $E^{\mathbb{N}}$ qui peut être muni d'une structure d'espace vectoriel normé pour la norme $\|\cdot\|_\infty$ définie par :

$$\forall\, u \in \ell^\infty(E),\ \|u\|_\infty = \sup_{n\in\mathbb{N}} \|u_n\|\,.$$

- **Suites convergentes**

  Soit $u \in E^{\mathbb{N}}$ une suite. On dit que u est convergente s'il existe $\ell \in E$ tel que :

  $$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N} \text{ tel que } \forall n \in \mathbb{N},\ n \geqslant n_0 \implies \|u_n - \ell\| < \varepsilon\,.$$

  Cela signifie que, pour toute boule $\mathcal{B}$ de centre ℓ et de rayon ε strictement positif (arbitrairement petit), la suite u est à valeurs dans $\mathcal{B}$ à partir d'un certain rang.
  Si u est convergente, alors le vecteur ℓ introduit dans la définition est unique. On dit alors que u converge vers ℓ, ou que ℓ est la limite de la suite u, et l'on note :

  $$\ell = \lim_{n \to +\infty} u_n\,.$$

  Une suite qui n'est pas convergente est dite divergente.

- **Propriétés des suites convergentes**

  – Dire que $\ell = \lim\limits_{n \to +\infty} u_n$ signifie aussi que la suite réelle $\big(\|u_n - \ell\|\big)$ converge vers 0 quand $n \to +\infty$.

  – Toute suite convergente est bornée.

  – Soient (u_n) et (v_n) deux suites d'éléments de E convergeant respectivement vers ℓ et ℓ'. Alors la suite $(u_n + v_n)$ converge vers $\ell + \ell'$.

  – Soit (λ_n) une suite d'éléments de $\mathbb{K}$ convergeant vers $\lambda \in \mathbb{K}$, et (u_n) une suite d'éléments de E convergeant vers $\ell \in E$. Alors, la suite $(\lambda_n u_n)$ converge vers $\lambda\ell$.

  Cela entraîne que l'ensemble des suites convergentes à valeurs dans E est un sous-espace vectoriel de $E^{\mathbb{N}}$. De plus, l'application qui à une suite convergente associe sa limite est linéaire.

- **Suites extraites**

  . Soit $u \in E^{\mathbb{N}}$. On appelle suite extraite de u toute suite de la forme $\big(u_{\varphi(n)}\big)$, où $\varphi\colon \mathbb{N} \to \mathbb{N}$ est une application strictement croissante.

  . Un élément $a \in E$ est appelé valeur d'adhérence d'une suite $u \in E^{\mathbb{N}}$ s'il existe une suite extraite de u qui converge vers a.

  . Si $u \in E^{\mathbb{N}}$ est une suite convergente de limite $\ell \in E$, toute suite extraite de u converge vers ℓ.

  Il en résulte qu'une suite qui a au moins deux valeurs d'adhérence distinctes est divergente.

- **Comparaison de normes**

  Soit E un $\mathbb{K}$-espace vectoriel, et soient N_1 et N_2 deux normes sur E.

  – Les deux normes N_1 et N_2 sont équivalentes si et seulement si toute suite $u \in E^{\mathbb{N}}$ bornée pour l'une est aussi bornée pour l'autre.

  – Les deux normes N_1 et N_2 sont équivalentes si et seulement si toute suite $u \in E^{\mathbb{N}}$ qui converge vers $\ell \in E$ au sens de l'une converge aussi vers ℓ au sens de l'autre.

  ✓ Ces résultats peuvent tomber en défaut si les normes ne sont pas équivalentes.

Exemple : dans $E = \mathscr{C}([0\,;1],\mathbb{R})$ muni des normes usuelles $\|\cdot\|_1$ et $\|\cdot\|_\infty$,

- la suite $f_n : t \mapsto nt^n$ est bornée pour $\|\cdot\|_1$, mais pas pour $\|\cdot\|_\infty$;
- la suite $g_n : t \mapsto t^n$ converge vers 0 pour $\|\cdot\|_1$, mais diverge pour $\|\cdot\|_\infty$ (plus précisément, cette suite converge simplement mais pas uniformément, comme nous le verrons dans le chapitre consacré aux suites de fonctions).

- **Cas d'un espace vectoriel normé de dimension finie**

Soit $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie p, muni d'une base $\mathscr{B} = (e_1, \ldots, e_p)$. Soit $u \in E^{\mathbb{N}}$, et soit $\ell \in E$. On note $(\ell_1, \ldots, \ell_p) \in \mathbb{K}^p$ les coordonnées de ℓ dans la base $\mathscr{B}$ et, pour tout $n \in \mathbb{N}$, on note $\left(u_n^{(1)}, \ldots, u_n^{(p)}\right) \in \mathbb{K}^p$ les coordonnées de u_n dans la base $\mathscr{B}$, c'est-à-dire :

$$\ell = \sum_{i=1}^{p} \ell_i e_i \qquad \text{et} \qquad \forall\, n \in \mathbb{N},\ u_n = \sum_{i=1}^{p} u_n^{(i)} e_i\,.$$

Les suites scalaires $(u_n^{(i)})$ ainsi définies sont appelées les suites coordonnées de u dans la base $\mathscr{B}$. Alors :

$$\lim_{n\to+\infty} u_n = \ell \iff \forall\, i \in [\![1\,;p]\!],\ \lim_{n\to+\infty} u_n^{(i)} = \ell_i\,.$$

Par conséquent, l'étude de la convergence d'une suite sur un espace vectoriel normé de dimension finie peut se ramener à l'étude de la convergence de ses suites coordonnées dans une base.

- **Théorème de Bolzano-Weierstrass**

Toute suite bornée d'un espace vectoriel normé *de dimension finie* possède une valeur d'adhérence.

> ✓ Ce résultat tombe en défaut si E n'est pas de dimension finie : considérer le $\mathbb{C}$-espace vectoriel $E = \mathscr{C}([0\,;2\pi],\mathbb{C})$ muni de la norme $\|\cdot\|_\infty$ et la suite de fonctions $f_n : t \mapsto \mathrm{e}^{int}$ (exemple détaillé page 380).

[S7.5] Topologie d'un espace vectoriel normé

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

- **Ouverts**

Une partie Ω de E s'appelle un ouvert si :

- soit : $\Omega = \emptyset$;
- soit : pour tout $a \in \Omega$, il existe $r > 0$ tel que $\mathcal{B}(a,r) \subset \Omega$.

Propriétés

- $\emptyset$ et E sont des ouverts.
- Une boule ouverte est un ouvert.
- La réunion d'une famille quelconque d'ouverts est un ouvert.
- L'intersection d'une famille *finie* d'ouverts est un ouvert.

- **Voisinages**

  Soit $a \in E$. On dit qu'une partie V de E est un voisinage de a s'il existe $r > 0$ tel que $\mathcal{B}(a,r) \subset V$.

  On note $\mathcal{V}(a)$ l'ensemble de voisinages de a.

  Une partie Ω de E est un ouvert si et seulement si Ω est voisinage de tous ses points.

  Dans le cas particulier $E = \mathbb{R}$, on définit les voisinages de $+\infty$ (respectivement $-\infty$) comme les parties de $\mathbb{R}$ contenant un intervalle de la forme $]a\,;+\infty[$ (respectivement $]-\infty\,;a[$).

- **Fermés**

  Soit $A \subset E$. On dit que A est un fermé de E (ou une partie fermée de E) si le complémentaire $\complement_E A$ de A dans E est un ouvert.

  Propriétés

  – $\emptyset$ et E sont des fermés (on peut démontrer que ce sont les seules parties à la fois ouvertes et fermées de E).

  – Une boule fermée, une sphère, sont des fermés.

  – L'intersection d'une famille quelconque de fermés est un fermé.

  – La réunion d'une famille finie de fermés est un fermé.

  – Un sous-espace vectoriel *de dimension finie* d'un espace vectoriel normé est un fermé.

- **Points intérieurs à une partie, intérieur**

  Soit A une partie non vide de E. On dit que $a \in A$ est intérieur à A s'il existe une boule ouverte non vide de centre a qui est incluse dans A.

  L'ensemble des points intérieurs à A s'appelle l'intérieur de A, et se note $\overset{\circ}{A}$:

  $$\overset{\circ}{A} = \{a \in A \mid \exists r > 0,\, \mathcal{B}(a,r) \subset A\}\,.$$

  Propriétés

  – $a \in \overset{\circ}{A}$ si et seulement si A est un voisinage de a.

  – $\overset{\circ}{A} \subset A$ et $\overset{\circ}{A} = A \iff A$ est un ouvert.

  – $\overset{\circ}{A}$ est le plus grand ouvert inclus dans A.

  – L'intérieur d'une boule fermée est la boule ouverte de même centre et de même rayon.

- **Points adhérents à une partie, adhérence**

  Soit A une partie non vide de E. On dit que $x \in E$ est un point adhérent à A si toute boule ouverte non vide de centre x contient au moins un élément de A.
  L'ensemble des points adhérents à A s'appelle l'adhérence de A et se note $\overline{A}$:

  $$x \in \overline{A} \iff \forall r > 0,\, \mathcal{B}(x,r) \cap A \neq \emptyset \iff \forall r > 0,\, \exists a \in A,\, d(x,a) < r\,.$$

  Propriétés

  – $x \in \overline{A} \iff d(x,A) = 0\,.$

- $A \subset \overline{A}$ et $A = \overline{A} \iff A$ est un fermé.
- $\overline{A}$ est le plus petit fermé contenant A.
- $\complement_E \overline{A} = \overset{\circ}{\complement_E A}$ et $\complement_E \overset{\circ}{A} = \overline{\complement_E A}$.
- L'adhérence d'une boule ouverte est la boule fermée de même centre et de même rayon.

- **Caractérisation séquentielle des points adhérents à une partie**

  Soit $A \subset E$, et soit $x \in E$. Alors x est un point adhérent à A si et seulement si il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers x.

  L'adhérence de A est donc l'ensemble des limites des suites convergentes d'éléments de A.

- **Caractérisation séquentielle d'un fermé**

  Une partie A de E est un fermé si et seulement si toute suite d'éléments de A qui converge dans E converge dans A.

- **Partie dense**

  Soit A une partie d'un espace vectoriel normé E. On dit que A est dense dans E si $\overline{A} = E$.

  Cela équivaut à dire que tout élément de E est limite d'une suite d'éléments de A, ou encore que toute boule de E rencontre A.

  Exemples

  - $\mathbb{Q}$ et $\complement_{\mathbb{R}}\mathbb{Q}$ sont denses dans $\mathbb{R}$.

    Cela signifie que tout réel est limite d'une suite de nombres rationnels ou d'une suite de nombres irrationnels ; ce résultat a été démontré en 1re année.
  - L'ensemble $\mathrm{GL}_n(\mathbb{K})$ des matrices carrées inversibles d'ordre n est dense dans $\mathcal{M}_n(\mathbb{K})$. Ce résultat important est démontré en **[SF7.34]** page 394.

    ✓ Pour ces exemples, il n'est pas utile de préciser la norme utilisée, puisque toutes les normes sur un espace vectoriel de dimension finie sont équivalentes. Cependant, pour les démonstrations, certaines normes peuvent être plus pratiques que d'autres.

- **Frontière**

  Soit $A \subset E$. On appelle frontière de A l'ensemble $\partial A = \overline{A} \setminus \overset{\circ}{A}$. La frontière de A est donc l'ensemble des points de E adhérents à A et non intérieurs à A.

  On peut aussi définir la frontière de A par : $\partial A = \overline{A} \cap \overline{\complement_E A}$.

  La frontière de A est donc l'ensemble des points adhérents à la fois à A et au complémentaire de A.

- **Normes équivalentes**

  Toutes les notions topologiques définies ci-dessus dépendent a priori du choix initial de la norme $\|\cdot\|$ sur E.

  Cependant, si deux normes N_1 et N_2 sont équivalentes, les parties ouvertes pour l'une le sont aussi pour l'autre. Il en résulte que toutes les autres notions (partie fermée, intérieur, adhérence etc...) sont aussi les mêmes pour les deux normes.

  En particulier, si E est de dimension finie, l'équivalence des normes assure que la topologie est indépendante de la norme choisie.

- **Topologie induite**
  Soit A une partie non vide d'un espace vectoriel normé E.
  - Soit $a \in A$. On appelle voisinage de a relatif à A l'intersection avec A d'un voisinage de a dans E.
  - On appelle ouvert relatif (respectivement fermé relatif) de A l'intersection avec A d'un ouvert de E (respectivement d'un fermé de E).
  - **Caractérisation séquentielle d'un fermé relatif**
    Une partie F de A est un fermé relatif à A si et seulement si toute suite d'éléments de F qui converge dans A converge dans F.

[S7.6] Compacité

- Une partie K d'un espace vectoriel normé $(E, \|\cdot\|)$ est dite compacte si de toute suite d'éléments de K on peut extraire une suite convergente vers un élément de K , autrement dit si toute suite d'éléments de K admet une valeur d'adhérence dans K.
- **Propriétés**
  - Si K est une partie compacte, alors K est une partie fermée et bornée.
  - Si K est un compact, et si F est un fermé inclus dans K, alors F est compact.
  - Toute suite d'éléments d'une partie compacte K qui admet une seule valeur d'adhérence ℓ est convergente dans K, de limite ℓ (ce résultat est démontré en **[SF7.5]** page 372).
  - Un produit d'une famille finie de compacts est un compact (dans l'espace vectoriel normé produit).
- Si E est un espace vectoriel normé *de dimension finie*, une partie de E est compacte si et seulement si elle est fermée et bornée.

  ✓ Un exemple de partie fermée bornée non compacte est donné en **[SF7.15]** page 380 .

[S7.7] Limites

$(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ sont ici deux espaces vectoriels normés, et A désigne une partie non vide de E.

- Soit f une application de A dans F, et soit $a \in \overline{A}$ un point adhérent à A.
  On dit que f admet une limite en a s'il existe $\ell \in F$ tel que :

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in A,\ \|x - a\|_E \leqslant \alpha \implies \|f(x) - \ell)\|_F \leqslant \varepsilon$$

  ou bien, en termes de voisinages :

$$\forall V \in \mathcal{V}(\ell),\ \exists U \in \mathcal{V}(a) \text{ tel que } f(U \cap A) \subset V\,.$$

  L'intérêt de la définition faisant intervenir des voisinages est qu'elle s'applique aussi au cas où $a = \pm\infty$ lorsque $E = \mathbb{R}$ et au cas $\ell = \pm\infty$ lorsque $F = \mathbb{R}$.
  Si f admet une limite en a, alors le vecteur ℓ introduit dans la définition est unique. On dit alors que ℓ est la limite de f en a, et on note :

$$\ell = \lim_{x \to a} f(x)\,.$$

- **Caractérisation séquentielle de la limite**
  Soit f une application de A dans F, soit $a \in \overline{A}$ et soit $\ell \in F$.
  Les propriétés suivantes sont équivalentes :
  (i) $\lim\limits_{x\to a} f(x) = \ell$,
  (ii) pour toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers a, la suite $\big(f(x_n)\big)_{n\in\mathbb{N}}$ converge vers ℓ.
- **Opérations algébriques sur les limites**
  Soit f et g deux applications de A dans F et φ une application de A dans $\mathbb{K}$. Soit $a \in \overline{A}$.
  – Si $\lim\limits_{x\to a} f(x) = \ell$ et $\lim\limits_{x\to a} g(x) = \ell'$ existent, alors $\lim\limits_{x\to a}(f+g)(x)$ existe et est égale à $\ell + \ell'$.
  – Si $\lim\limits_{x\to a} f(x) = \ell$ et $\lim\limits_{x\to a} \varphi(x) = \lambda\,(\in \mathbb{K})$ existent, alors $\lim\limits_{x\to a}(\varphi.f)(x)$ existe et est égale à $\lambda\ell$.
- **Limite d'une composée**
  Soient E, F et G trois espaces vectoriels normés, $f\colon A \subset E \to F$ et $g\colon B \subset F \to G$ deux applications.
  On suppose $f(A) \subset B$, et $\lim\limits_{x\to a} f(x) = b$.
  Alors $b \in \overline{B}$, et si $\lim\limits_{y\to b} g(y) = \ell \in G$, alors $\lim\limits_{x\to a}(g\circ f)(x) = \ell$.
- **Cas de la dimension finie**
  On suppose ici que f est une application définie sur une partie A d'un espace vectoriel normé E, à valeurs dans un espace vectoriel normé F de *dimension finie*. Soit $\mathscr{B} = (e_1,\dots,e_p)$ une base de F.
  On note $(f_1,\dots,f_p)$ les applications coordonnées de f dans la base $\mathscr{B}$, c'est-à-dire :
  $$\forall\, x \in A,\ f(x) = \sum_{i=1}^{p} f_i(x)e_i \quad \text{avec} \quad f_i : A \to \mathbb{K}\,.$$
  Soit $a \in \overline{A}$. Pour que f admette une limite $\ell = \sum\limits_{i=1}^{p} \ell_i e_i \in F$ quand x tend vers a, il faut et il suffit que, pour tout $i \in [\![1\,;p]\!]$, f_i admette une limite dans $\mathbb{K}$ quand x tend vers a, et on a alors : $\lim\limits_{x\to a} f_i(x) = \ell_i$.

[S7.8] Continuité

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, et D une partie non vide de E.

- Soit f une application de D dans F, et soit a *appartenant à* D. On dit que f est continue en a si f admet une limite en a (qui vaut alors nécessairement $f(a)$), c'est-à-dire :
  $$\forall\,\varepsilon > 0,\ \exists\,\alpha > 0,\ \forall\, x \in D,\ \|x-a\|_E \leqslant \alpha \implies \|f(x)-f(a)\|_F \leqslant \varepsilon\,.$$
  On dit que f est continue sur D si f est continue en tout point de D. On note $\mathscr{C}(D,F)$ l'ensemble des applications continues sur D à valeurs dans F.

- Les notions de limite et de continuité sont inchangées si l'on remplace les normes par des normes équivalentes. Lorsque E et F sont de dimensions finies, ces notions sont donc indépendantes des normes choisies.
  En revanche, si deux normes ne sont pas équivalentes, une application peut être continue au sens d'une norme et pas de l'autre.

- **Caractérisation séquentielle de la continuité**
  Soit f une application de D dans F, et $a \in D$.
  f est continue en a si et seulement si pour toute suite (x_n) d'éléments de D qui converge vers a, la suite $\big(f(x_n)\big)$ converge vers $f(a)$.

- **Opérations algébriques**
  Soient f et g deux applications de D dans F et φ une application de D dans $\mathbb{K}$. Soit $a \in D$.
  - Si f et g sont continues en a, alors $f+g$ est continue en a.
  - Si f et φ sont continues en a, l'application $\varphi.f$ est continue en a.

  En particulier, $\mathscr{C}(D,F)$ est un sous-espace vectoriel de $\mathcal{A}(D,F)$ et $\mathscr{C}(D,\mathbb{K})$ est une sous-algèbre de $\mathcal{A}(D,\mathbb{K})$.

- **Composée d'applications continues**
  Soient E, F et G trois espaces vectoriels normés, D une partie de E et Δ une partie de F.
  Soit f une application de D dans F telle que $f(D) \subset \Delta$, et g une application de Δ dans G.
  - Si f est continue en $a \in A$ et si g est continue en $f(a)$, $g \circ f$ est continue en a.
  - Si $f \in \mathscr{C}(D,\Delta)$ et $g \in \mathscr{C}(\Delta,G)$, alors $g \circ f \in \mathscr{C}(D,G)$.

- **Unicité du prolongement par continuité**
  Soient E et F deux espaces vectoriels normés et D une partie de E. Soient f et g deux applications continues de D dans F. Soit A une partie dense dans D (c'est-à-dire $\overline{A} = D$).
  Si f et g coïncident sur A, alors $f = g$.

- **Image réciproque d'un ouvert ou d'un fermé**
  Soient E et F deux espaces vectoriels normés, et f une application d'une partie D de E dans F. Les propositions suivantes sont équivalentes :
  (i) f est continue sur D,
  (ii) l'image réciproque par f de tout ouvert de F est un ouvert relatif de D,
  (iii) l'image réciproque par f de tout fermé de F est un fermé relatif de D.

  Exemples

  . Soient $f, g\colon \mathbb{R} \longrightarrow \mathbb{R}$, continues. Alors :

  $$\{x \in \mathbb{R} \mid f(x) < g(x)\} \text{ est un ouvert de } \mathbb{R}.$$
  $$\{x \in \mathbb{R} \mid f(x) \leqslant g(x)\} \text{ est un fermé de } \mathbb{R}.$$

  (En effet, le premier ensemble est l'image réciproque par l'application continue $f-g$ de l'ouvert $]-\infty\,;0[$ et le second est l'image réciproque du fermé $]-\infty\,;0]$).

- Soit $f \in \mathscr{C}(\mathbb{R},\mathbb{R})$ et soient

$$\Gamma = \{(x,y) \in \mathbb{R}^2 \mid y = f(x)\} \text{ le graphe de } f\,,$$
$$\mathscr{E} = \{(x,y) \in \mathbb{R}^2 \mid y > f(x)\} \text{ l'épigraphe strict de } f\,.$$

  Alors Γ est une partie fermée de $\mathbb{R}^2$ et $\mathscr{E}$ est une partie ouverte de $\mathbb{R}^2$.

  (En effet, il suffit ici de considérer l'application continue $(x,y) \mapsto y - f(x)$ de $\mathbb{R}^2$ dans $\mathbb{R}$. Γ est l'image réciproque du fermé $\{0\}$, et $\mathscr{E}$ celle de l'ouvert $]0\,;+\infty[$).

- L'ensemble $\mathrm{GL}_n(\mathbb{K})$ des matrices carrées inversibles d'ordre n est un ouvert de $\mathcal{M}_n(\mathbb{K})$.

  (En effet, c'est l'image réciproque de l'ouvert $\mathbb{K} \setminus \{0\}$ par l'application « déterminant », qui est continue de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$).

- **Cas de la dimension finie**

  On suppose ici que F est de *dimension finie*. On munit F d'une base $\mathscr{B} = (e_1,\ldots,e_p)$. Soit $f \in \mathcal{A}(D,F)$, et $(f_1,\ldots,f_p)$ les applications coordonnées de f dans la base $\mathscr{B}$. Alors :

  f est continue en $a \in D$ si et seulement si pour tout $i \in [\![1\,;n]\!]$ f_i est continue en a.

- **Continuité uniforme**

  - Soient E et F deux espaces vectoriels normés, D une partie non vide de E, f une application de D dans F.

    On dira que f est uniformément continue sur D si et seulement si :

$$\forall\,\varepsilon > 0,\ \exists\,\alpha > 0,\ \forall\,(x,y) \in D^2,\ \|x-y\|_E < \alpha \Longrightarrow \|f(x)-f(y)\|_F < \varepsilon\,.$$

    ✓ La différence entre cette notion (qui est globale) et la notion de continuité en un point (notion locale) est que, dans la propriété ci-dessus, α ne dépend *que de* ε, et non pas de x ou y.

  - Si f est uniformément continue sur D, alors f est continue sur D.

    ✓ La réciproque de cette propriété est fausse : il existe des fonctions continues qui ne sont pas uniformément continues, comme le montre l'exemple de $x \mapsto x^2$ sur $\mathbb{R}$.

  - **Théorème de Heine**

    Soient E et F deux espaces vectoriels normés, K un compact de E, et $f\colon K \to F$ une application continue sur K.

    Alors f est uniformément continue sur K.

- **Fonctions lipschitziennes**

  - Soit f une application de D dans F. On dit que f est lipschitzienne sur D s'il existe un réel k (positif) tel que :

$$\forall\,(x,y) \in E^2,\ \|f(x)-f(y)\|_F \leqslant k\,\|x-y\|_E\,.$$

    On dit alors que f est lipschitzienne de rapport k.

- Si f est lipschitzienne sur D, alors f est uniformément continue sur D.

  ✓ La réciproque de cette propriété est fausse : par exemple, la fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}_+$ (cela est démontré en **[SF7.30]** page 391), sans être lipschitzienne sur $\mathbb{R}_+$.

- **Continuité et compacité**
  - Soit K une partie compacte d'un espace vectoriel normé, et f une application continue de K dans un espace vectoriel normé F.
    Alors $f(K)$ est une partie compacte de F.
  - Cela entraîne que si f est continue sur K, alors f est bornée et $\|f\|$ admet sur K un maximum et un minimum.
  - Soit f une application continue d'un compact A de E à valeurs dans F, et soit $B = f(A)$.
    Si f est bijective de A sur B, alors f^{-1} est continue sur B.

- **Caractérisation des applications linéaires continues**
  Soient E et F deux espaces vectoriels normés, et $u \in \mathscr{L}(E,F)$. Les propriétés suivantes sont équivalentes :
  (i) u est continue,
  (ii) u est continue en 0_E,
  (iii) u est bornée sur la boule unité fermée $\mathcal{B}_f(0,1) = \{x \in E \mid \|x\|_E \leqslant 1\}$,
  (iv) il existe un réel $k \geqslant 0$ tel que pour tout $x \in E$, $\|u(x)\|_F \leqslant k\,\|x\|_E$,
  (v) u est lipschitzienne,
  (vi) u est uniformément continue.

- **Caractérisation des applications multilinéaires continues**
  La définition d'une application p-linéaire a été donnée en **[S3.14]**.
  Soient $(E_i)_{1\leqslant i\leqslant p}$ p espaces vectoriels normés, muni chacun d'une norme notée $\|\cdot\|_i$.
  On munit l'espace vectoriel produit $E = \prod_{i=1}^{p} E_i$ de la norme $\|\cdot\|_\infty$ définie par :

$$\forall\, x = (x_1,\ldots,x_p) \in E, \quad \|x\|_\infty = \max_{1\leqslant i\leqslant p} \|x_i\|_i \ .$$

  Soit F un espace vectoriel normé, et $u\colon E \to F$ une application p-linéaire. Alors, les propriétés suivantes sont équivalentes :
  (i) u est continue,
  (ii) u est continue en 0_E,
  (iii) u est bornée sur le produit $B_1 \times \ldots \times B_p$, où B_i est la boule unité fermée de E_i : $B_i = \{x_i \in E_i \mid \|x_i\|_i \leqslant 1\}$,
  (iv) il existe un réel $k \geqslant 0$ tel que, pour tout $x = (x_1,\ldots,x_p) \in E$, on ait

$$\|u(x_1,\ldots,x_p)\|_F \leqslant k\,\|x_1\|_1 \cdot \|x_2\|_2 \cdots \|x_p\|_p \ .$$

- **Exemples fondamentaux d'applications continues**
  - Si E est de *dimension finie*, alors toute application linéaire $u\colon E \to F$ est continue sur E.
  - Si $E_1, \dots, E_p$ sont des espaces vectoriels normés *de dimensions finies*, toute application p-linéaire sur $E_1 \times \cdots \times E_p$ est continue.
  - Toute application polynomiale $\varphi\colon \mathbb{K}^n \to \mathbb{K}$ est continue.
  - L'application det$\colon \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est continue.

[S7.9] Parties connexes par arcs

- Soit E un espace vectoriel normé, et a, $b \in E$. Un chemin continu reliant a à b est une application $\gamma\colon \left\{ \begin{array}{ccc} [0\,;1] & \longrightarrow & E \\ t & \longmapsto & \gamma(t) \end{array} \right.$, continue et telle que $\gamma(0) = a$ et $\gamma(1) = b$.
  On appelle support de γ l'ensemble image $\gamma([0\,;1]) = \{\gamma(t),\, t \in [0\,;1]\}$.
- Soit A une partie non vide de E. La relation binaire $\mathcal{R}$ définie sur A^2 par :
  $$a\mathcal{R}b \iff \begin{array}{l} \text{il existe un chemin continu reliant } a \text{ à } b \\ \text{et dont le support est inclus dans } A \end{array}$$
  est une relation d'équivalence.
  Ses classes d'équivalence s'appellent les composantes connexes par arcs de A.
- A est dite connexe par arcs si elle possède une seule composante connexe, c'est-à-dire si pour tous a, $b \in A$ il existe un chemin continu reliant a à b et dont le support est inclus dans A.
  Exemples
  - Toute partie convexe de E est connexe par arcs.
  - Une partie A est dite étoilée par rapport à un de ses points ω si pour tout $a \in A$ le segment $[\omega\,;a]$ est inclus dans A (une partie convexe est donc une partie qui est étoilée par rapport à tous ses points).
    Si A est étoilée par rapport à l'un de ses points, elle est connexe par arcs.
  - Les parties connexes par arcs de $\mathbb{R}$ sont les intervalles.
- **Image par une application continue**
  Soient E et F deux espaces vectoriels normés, A une partie non vide de E et $f\colon A \to F$ une application continue.
  Si A est une partie connexe par arcs de E, son image $f(A)$ est une partie connexe par arcs de F.
- **Théorème des valeurs intermédiaires**
  Soit E un espace vectoriel normé, A une partie connexe par arcs de E et f une application continue de A dans $\mathbb{R}$.
  Alors $f(A)$ est un intervalle de $\mathbb{R}$; autrement dit, si f prend deux valeurs réelles $x \leqslant y$ sur A, pour tout $t \in [x\,;y]$ il existe $a \in A$ tel que $f(a) = t$.

Thème 8 - Suites numériques

Les suites numériques ont été étudiées en première année, et nous avons vu dans le chapitre précédent les principales définitions et résultats concernant plus généralement les suites à valeurs vectorielles.
Cependant il nous a semblé important d'introduire un bref chapitre de révisions à ce sujet compte tenu de l'importance qu'ont les suites numériques dans tout le reste du programme. De plus, il existe des résultats spécifiques aux suites à valeurs réelles, liés à la relation d'ordre sur $\mathbb{R}$.

[S8.1] Limite d'une suite réelle

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles.

- On dit que la suite u admet le réel ℓ pour limite si :

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |u_n - \ell| < \varepsilon\,.$$

- On dit que la suite u tend vers $+\infty$ (respectivement $-\infty$) si :

$$\forall A > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ u_n > A \text{ (respectivement } u_n < -A)\,.$$

✓ Dire qu'une suite $(u_n)_{n\in\mathbb{N}}$ est convergente signifie que $\lim\limits_{n\to+\infty} u_n = \ell$ où ℓ est un réel fini. Une suite qui tend vers $\pm\infty$ est divergente.

[S8.2] Composition des limites

Soit f une fonction numérique et $(u_n)_{n\in\mathbb{N}}$ une suite réelle telle que u_n appartienne au domaine de définition de f au moins à partir d'un certain rang.

- Si $\lim\limits_{n\to+\infty} u_n = a$ et si $\lim\limits_{x\to a} f(x) = \ell$, alors $\lim\limits_{n\to+\infty} f(u_n) = \ell$.
- Si $\lim\limits_{n\to+\infty} u_n = a$ et si f est continue en a, alors $\lim\limits_{n\to+\infty} f(u_n) = f(a)$.

[S8.3] Propriétés des suites réelles liées à l'ordre

- **Prolongement des inégalités (cas de la convergence)**
  Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles convergeant respectivement vers ℓ et ℓ'.
  - Si $\ell < \ell'$, il existe un entier n_0 tel que $\forall n \in \mathbb{N},\ n \geqslant n_0 \implies u_n < v_n$.
  - S'il existe un entier n_0 tel que $\forall n \in \mathbb{N},\ n \geqslant n_0 \implies u_n \leqslant v_n$, alors $\ell \leqslant \ell'$.

    ✓ Pour appliquer cette propriété, il faut préalablement démontrer *l'existence* des limites.

    ✓ Si pour tout $n \geqslant n_0$ on a $u_n < v_n$, on peut quand même avoir $\ell = \ell'$. Le passage à la limite affaiblit la relation d'ordre.

- **Prolongement des inégalités (cas de la divergence)**
  Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles telles que :

$$\exists n_0 \in \mathbb{N} \text{ tel que } \forall n \in \mathbb{N},\ n \geqslant n_0 \implies u_n \leqslant v_n\,.$$

  Alors on a :

$$\lim_{n\to+\infty} u_n = +\infty \implies \lim_{n\to+\infty} v_n = +\infty$$

  et

$$\lim_{n\to+\infty} v_n = -\infty \implies \lim_{n\to+\infty} u_n = -\infty\,.$$

- **Théorème d'encadrement**
  Soit $(u_n)_{n\in\mathbb{N}}$, $(v_n)_{n\in\mathbb{N}}$ et $(w_n)_{n\in\mathbb{N}}$ trois suites réelles, telles que :

$$\exists n_0 \in \mathbb{N} \text{ tel que } \forall n \in \mathbb{N},\ n \geqslant n_0 \implies u_n \leqslant v_n \leqslant w_n$$

  et

$$\lim_{n\to+\infty} u_n = \lim_{n\to+\infty} w_n = \ell\,.$$

  Alors la suite $(v_n)_{n\in\mathbb{N}}$ est convergente, et $\lim\limits_{n\to+\infty} v_n = \ell$.

- **Théorème de la limite monotone**
  Soit (u_n) une suite de réels croissante (respectivement décroissante), au moins à partir d'un certain rang.
  Alors (u_n) est convergente si et seulement si elle est majorée (respectivement minorée).
  (On peut préciser : si (u_n) est croissante non majorée, alors $\lim\limits_{n\to+\infty} u_n = +\infty$).

[S8.4] Suites adjacentes

- **Définition**
  Soit$(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites réelles.
  On dit que ces deux suites sont adjacentes si :

$$\begin{cases} (u_n) \text{ est croissante} \\ (v_n) \text{ est décroissante} \\ \lim\limits_{n\to+\infty} (v_n - u_n) = 0 \end{cases}.$$

- **Théorème**
  Deux suites adjacentes sont convergentes, vers la même limite.

- On a de plus : $\forall\,(p,q) \in \mathbb{N}^2$, $u_p \leqslant \ell \leqslant v_q$ (avec inégalités strictes si les deux suites sont strictement monotones).
  Cette remarque est importante car elle permet d'obtenir facilement un encadrement de ℓ.

[S8.5] Suites arithmético-géométriques

On considère ici l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ vérifiant une relation de récurrence de la forme :

$$\forall\, n\in\mathbb{N},\ u_{n+1}=au_n+b$$

avec $a,\ b\in\mathbb{K}$.

- Si $a=1$, la suite est tout simplement arithmétique de raison b et l'on a alors $u_n=u_0+nb$ pour tout entier n.
- Si $a\neq 1$, la suite constante égale à $\ell=\dfrac{b}{1-a}$ vérifie bien la relation de récurrence (il suffit de résoudre l'équation $\ell=a\ell+b$). Si l'on pose alors $v_n=u_n-\ell$ pour tout n, on a :

$$v_{n+1}=u_{n+1}-\ell=(au_n+b)-(a\ell+b)=a(u_n-\ell)=av_n\,,$$

de sorte que la suite $(v_n)_{n\in\mathbb{N}}$ est géométrique de raison a. On aura donc $v_n=a^nv_0$ pour tout n, d'où l'on tire finalement :

$$\forall\, n\in\mathbb{N},\ u_n=v_n+\ell=a^n\left(u_0-\frac{b}{1-a}\right)+\frac{b}{1-a}\cdot$$

✓ Il ne faut bien sûr pas retenir cette dernière formule par cœur mais seulement la méthode ; les calculs sont alors faciles à refaire.

✓ Il faut savoir adapter les résultats dans le cas où la suite n'est définie qu'à partir d'un rang n_0.

Pour cela, il suffit de se rappeler que les formules deviennent :

- $u_n=u_{n_0}+(n-n_0)b$ pour une suite arithmétique de raison b ;
- $v_n=a^{n-n_0}v_{n_0}$ pour une suite géométrique de raison a.

[S8.6] Suites récurrentes linéaires d'ordre 2

On considère ici l'ensemble $\mathscr{S}$ des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ qui vérifient une relation de la forme :

$$\forall\, n\in\mathbb{N},\ au_{n+2}+bu_{n+1}+cu_n=0\qquad (L)$$

avec $a,\ b,\ c\in\mathbb{K}$, $a\neq 0$.
Il est facile de vérifier que $\mathscr{S}$ est un sous-espace vectoriel de l'espace vectoriel $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans $\mathbb{K}$.
On montre également que l'application $\varphi:\left\{\begin{array}{ccc}\mathscr{S} & \longrightarrow & \mathbb{K}^2\\ u & \longmapsto & (u_0,u_1)\end{array}\right.$ qui à toute suite $u\in\mathscr{S}$ associe le couple formé par ses deux premiers termes, est linéaire et bijective.
φ est donc un isomorphisme d'espaces vectoriels et par conséquent, $\dim\mathscr{S}=2$.

On peut alors chercher des éléments particuliers de $\mathscr{S}$: une suite géométrique $n\mapsto r^n$ avec $r\in\mathbb{K}^*$ appartient à $\mathscr{S}$ si et seulement si

$$\forall\, n\in\mathbb{N},\ ar^{n+2}+br^{n+1}+cr^n=0\quad\text{soit}\quad ar^2+br+c=0\,.$$

L'équation (E_c) : $ar^2 + br + c = 0$ s'appelle l'équation caractéristique de la récurrence.
On étudie alors ses solutions. Il faut distinguer deux cas, selon que le corps de base est $\mathbb{R}$ ou $\mathbb{C}$. En notant Δ le discriminant de l'équation caractéristique, on a les résultats suivants.

- Si $\mathbb{K} = \mathbb{C}$
  - Si $\Delta \neq 0$, (E_c) possède deux racines distinctes r_1 et r_2 ; dans ce cas, les suites $n \mapsto r_1^n$ et $n \mapsto r_2^n$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est donc l'ensemble des suites de la forme :
  $$u_n = \lambda_1 r_1^n + \lambda_2 r_2^n \ , \ (\lambda_1,\lambda_2) \in \mathbb{C}^2 \, .$$
  - Si $\Delta = 0$, (E_c) possède une racine double r ; dans ce cas, les suites $n \mapsto r^n$ et $n \mapsto nr^n$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est alors l'ensemble des suites de la forme :
  $$u_n = (\lambda n + \mu) r^n \ , \ (\lambda,\mu) \in \mathbb{C}^2 \, .$$
- Si $\mathbb{K} = \mathbb{R}$
  - Si $\Delta \geqslant 0$, on a les mêmes résultats que ci-dessus (mais avec ici λ_1, λ_2, λ et μ réels).
  - Si $\Delta < 0$, l'équation (E_c) admet deux racines complexes non réelles conjuguées $re^{\pm i\theta}$, et alors les suites $n \mapsto r^n \cos n\theta$ et $n \mapsto r^n \sin n\theta$ sont deux solutions linéairement indépendantes de (L). L'ensemble des solutions de (L) est ici l'ensemble des suites de la forme :
  $$u_n = (\lambda \cos n\theta + \mu \sin n\theta) r^n \ , \ (\lambda,\mu) \in \mathbb{R}^2 \, .$$

  ✓ Les suites arithmético-géométriques ou les suites récurrentes linéaires peuvent apparaître dans plusieurs types d'exercices (calculs de déterminants, calculs de probabilités...). Il est donc important d'apprendre les méthodes rappelées ci-dessus.

Thème 9 - Fonctions réelles d'une variable réelle

Une fonction numérique d'une variable réelle est une application définie sur un intervalle I de $\mathbb{R}$, à valeurs dans $\mathbb{R}$. Avant d'aborder les fonctions à valeurs vectorielles dans le thème suivant, il nous a semblé important de rappeler certains résultats de 1re année ; ceux-ci sont complétés par la notion de fonction convexe.

[S9.1] Rappels du cours de 1re année

- **Théorème de la limite monotone**
  - Soient $]a\,;b[$ un intervalle ouvert (borné ou non) et f une application *croissante* de $]a\,;b[$ dans $\mathbb{R}$. Alors :
    - f admet en tout point x_0 de $]a\,;b[$ une limite à droite et une limite à gauche, notées respectivement $f(x_0^+)$ et $f(x_0^-)$, telles que $f(x_0^-) \leqslant f(x_0) \leqslant f(x_0^+)$.
    - f admet en b^- une limite, finie ou non ; cette limite est finie si et seulement si f est majorée, et dans le cas contraire, elle vaut $+\infty$.
    - f admet en a^+ une limite, finie ou non ; cette limite est finie si et seulement si f est minorée, et dans le cas contraire, elle vaut $-\infty$.
  - Dans le cas où f est *décroissante*, on a :
    - f admet en tout point x_0 de $]a\,;b[$ une limite à droite et une limite à gauche, telles que $f(x_0^-) \geqslant f(x_0) \geqslant f(x_0^+)$.
    - f admet en b^- une limite, finie ou non ; cette limite est finie si et seulement si f est minorée, et dans le cas contraire, elle vaut $-\infty$.
    - f admet en a^+ une limite, finie ou non ; cette limite est finie si et seulement si f est majorée, et dans le cas contraire, elle vaut $+\infty$.
- **Théorème de la bijection**

  Soit $f\colon I \longrightarrow \mathbb{R}$ une application *continue* et *strictement monotone* sur un intervalle I de $\mathbb{R}$.

  Alors $J = f(I)$ est un intervalle de $\mathbb{R}$, f est bijective de I sur J, et $f^{-1} : J \longrightarrow I$ est continue, strictement monotone, de même sens de variation que f.
- **Dérivée de la fonction réciproque**

  Soit $f\colon I \longrightarrow J$ une bijection continue de I sur J.

  Si f est dérivable en $a \in I$ et si $f'(a) \neq 0$, alors f^{-1} est dérivable en $b = f(a)$ et on a :
  $$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f' \circ f^{-1}(b)}\cdot$$
- **Caractérisation des $\mathscr{C}^k$-difféomorphismes**

  Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Soit f une application de classe $\mathscr{C}^k$ d'un intervalle I sur l'intervalle $J = f(I)$, telle que f' *ne s'annule pas sur* I.

  Alors f est bijective de I sur J, et f^{-1} est de classe $\mathscr{C}^k$ sur J.

  ✓ Une telle application s'appelle un $\mathscr{C}^k$-difféomorphisme de I sur J.

- Soit f une fonction numérique définie sur un intervalle I, et admettant en un point t_0 *intérieur à* I un extrémum relatif.
  Si f est dérivable en t_0, alors $f'(t_0) = 0$.

  ✓ Le fait que t_0 est un point intérieur est indispensable ! Par exemple, la fonction $f : \left\{ \begin{array}{ccc} [0\,;1] & \longrightarrow & [0\,;1] \\ t & \longmapsto & t \end{array} \right.$ admet un minimum en 0 et un maximum en 1, mais sa dérivée ne s'annule jamais !

- **Théorème de Rolle**
  Soit f une fonction numérique continue sur l'intervalle fermé $[a\,;b]$ $(a < b)$, dérivable sur l'intervalle ouvert $]a\,;b[$, et telle que $f(a) = f(b)$.
  Alors, il existe $c \in]a\,;b[$ tel que $f'(c) = 0$.

- **Théorème des accroissements finis**
  Soit f une fonction numérique continue sur l'intervalle fermé $[a\,;b]$ $(a < b)$ et dérivable sur l'intervalle ouvert $]a\,;b[$.
  Alors, il existe c appartenant à $]a\,;b[$ tel que : $f(b) - f(a) = (b-a)f'(c)$.
  Interprétation géométrique
  Si f est une fonction numérique continue sur $[a\,;b]$ et dérivable sur $]a\,;b[$, il existe (au moins) un point du graphe de f où la tangente est parallèle à la corde qui joint les points A de coordonnées $\big(a,f(a)\big)$ et B de coordonnées $\big(b,f(b)\big)$.

- **Théorème de prolongement de la dérivée**
  Soit $f : [a\,;b] \to \mathbb{R}$. On suppose que :
  - f est continue sur $[a\,;b]$;
  - f est dérivable sur $]a\,;b]$;
  - $\lim\limits_{x \to a^+} f'(x) = \ell \in \overline{\mathbb{R}}$.

  Alors : $\lim\limits_{x \to a^+} \dfrac{f(x) - f(a)}{x - a} = \ell$.
  Lorsque ℓ est finie, on en déduit que f est dérivable en a^+ et que $f'_d(a) = \ell$.
  Lorsque $\ell = \pm\infty$, la courbe de f admet au point de coordonnées $\big(a,f(a)\big)$ une tangente « verticale », et f n'est pas dérivable en a^+.

[S9.2] Convexité dans un $\mathbb{R}$-espace vectoriel

Dans ce paragraphe, E désigne un $\mathbb{R}$-espace vectoriel, dont les éléments seront indifféremment appelés points ou vecteurs. Si deux éléments a et b de E sont notés comme des points A et B, on définit le vecteur $\overrightarrow{AB}$ par $\overrightarrow{AB} = b - a$.

- **Barycentres**
  Soit $n \in \mathbb{N}^*$, soient $A_1, \ldots, A_n$ des points de E et $\lambda_1, \ldots, \lambda_n$ des réels tels que $\sum\limits_{i=1}^{n} \lambda_i \neq 0$. On appelle barycentre du système de points pondérés $(A_i,\lambda_i)_{1 \leqslant i \leqslant n}$ l'unique point G de E qui vérifie :

$$\forall\, O \in E,\ \overrightarrow{OG} = \frac{1}{\sum\limits_{i=1}^{n} \lambda_i} \sum_{i=1}^{n} \lambda_i \overrightarrow{OA_i}\,.$$

G est aussi l'unique point de E qui vérifie : $\sum_{i=1}^{n} \lambda_i \overrightarrow{GA_i} = \overrightarrow{0}$.

Propriétés

– Le barycentre d'un système de points pondérés $(A_i,\lambda_i)_{1\leqslant i\leqslant n}$ est inchangé si l'on multiplie tous les poids λ_i par un réel non nul. Dans la définition de G, on peut donc toujours faire en sorte d'avoir $\sum_{i=1}^{n} \lambda_i = 1$.

– Le barycentre d'un système de points pondérés $(A_i,\lambda_i)_{1\leqslant i\leqslant n}$ est inchangé si l'on remplace un sous-système de poids total non nul par son barycentre partiel affecté de ce poids total (c'est l'associativité du barycentre).

- **Caractérisation des parties convexes**

  . Soient A, B deux points de E. Rappelons que le segment $[A\,;B]$ est défini par :

  $$[A\,;B] = \{\lambda A + (1-\lambda)B,\ \lambda \in [0\,;1]\}\ .$$

  Il s'agit donc de l'ensemble des barycentres de A et B à coefficients positifs.

  . Rappelons également qu'une partie $\mathcal{C}$ de E est dite convexe si pour tous points A et B appartenant à $\mathcal{C}$, le segment $[A\,;B]$ est inclus dans $\mathcal{C}$.
  Puisque l'intersection de parties convexes de E, si elle est non vide, est encore une partie convexe de E, on peut définir l'enveloppe convexe d'une partie non vide $X \subset E$ comme étant l'intersection de toutes les parties convexes de E contenant X ; c'est donc le plus petit convexe de E (au sens de l'inclusion) qui contient X.

  . **Théorème**
  L'enveloppe convexe d'une famille de points $(A_i)_{1\leqslant i\leqslant n}$ est exactement l'ensemble des barycentres des A_i affectés de poids positifs.

  . Il en résulte qu'une partie $\mathcal{C}$ de E est convexe si et seulement si le barycentre de toute famille de points pondérés $(A_i,\lambda_i)_{1\leqslant i\leqslant n}$ avec $A_i \in \mathcal{C}$ et $\lambda_i \geqslant 0$ pour tout i, appartient encore à $\mathcal{C}$.

[S9.3] Fonctions numériques convexes

- Une fonction $f : I \to \mathbb{R}$, où I est un intervalle de $\mathbb{R}$, est dite convexe si :

  $$\forall\,(x,y) \in I^2,\ \forall\,\lambda \in [0\,;1],\ f\big(\lambda x + (1-\lambda)y\big) \leqslant \lambda f(x) + (1-\lambda) f(y)\,.$$

  f est dite strictement convexe si $f\big(\lambda x + (1-\lambda)y\big) < \lambda f(x) + (1-\lambda)f(y)$ dès que $x < y$ et $\lambda \in\,]0\,;1[$.
  f est dite concave si $-f$ est convexe (soit $f\big(\lambda x + (1-\lambda)y\big) \geqslant \lambda f(x) + (1-\lambda)f(y)$ pour tous x, $y \in I$ et tout $\lambda \in [0\,;1]$).

- **Inégalité de convexité (Jensen)**
  Soit $f : I \to \mathbb{R}$ une fonction convexe. Alors pour tout $n \in \mathbb{N}^*$, si $\lambda_1, \ldots, \lambda_n$ sont des réels positifs tels que $\sum_{i=1}^{n} \lambda_i = 1$ et si $(x_1, \ldots, x_n)$ sont des éléments de I, on a :

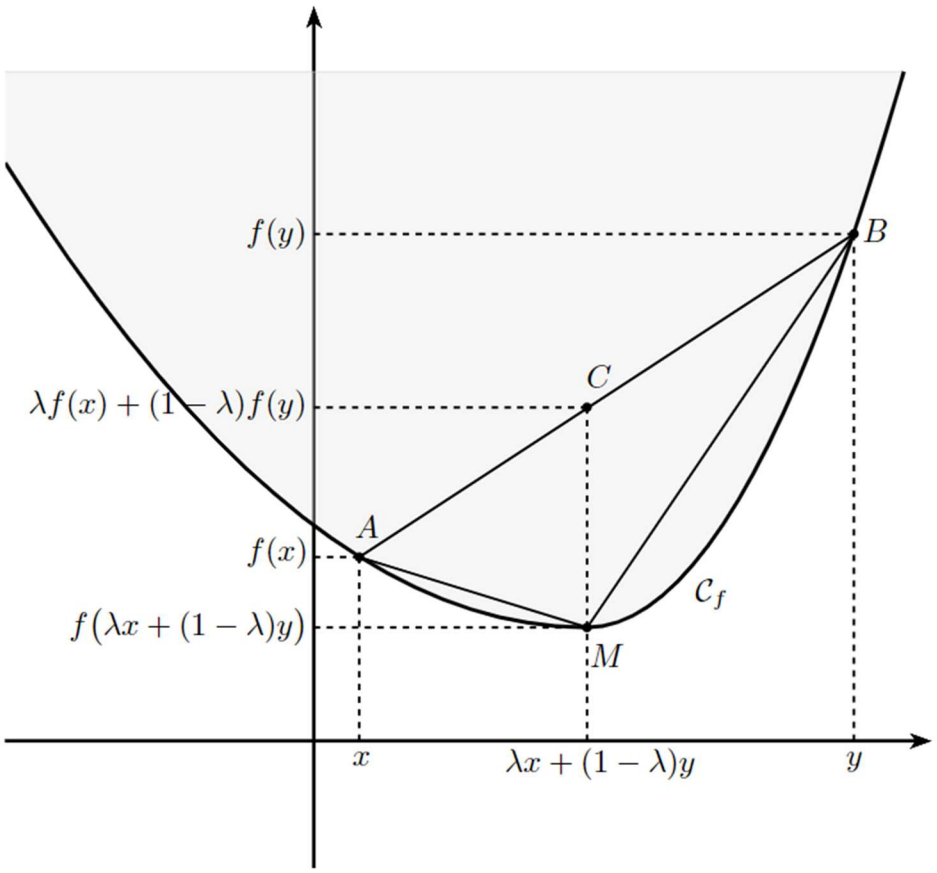
  $$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

(en abrégé : l'image du barycentre est inférieur au barycentre des images).
De plus, lorsque f est strictement convexe, il y a égalité si et seulement si tous les x_i sont égaux (lorsque $\lambda_i \in]0\,;1[$).

- **Interprétation géométrique**
  - Dire que $f\colon I \to \mathbb{R}$ est convexe signifie que, pour tous points A et B de la courbe $\mathcal{C}_f$ de f, l'arc $\overset{\frown}{AB}$ de la courbe est situé au-dessous du segment $[A\,;B]$ (sur la figure ci-après, le point M est au-dessous du point C).
  - On appelle épigraphe de f l'ensemble :

$$\mathcal{E}_f = \{(x,y) \in \mathbb{R}^2 \mid x \in I \text{ et } y \geqslant f(x)\}$$

(cela correspond donc à l'ensemble des points du plan situés au-dessus de la courbe de f, en gris sur la figure).
La fonction f est convexe si et seulement si $\mathcal{E}_f$ est une partie convexe de $\mathbb{R}^2$.



- **Inégalité des trois pentes**
  Soit $f\colon I \to \mathbb{R}$. f est convexe sur I si et seulement si :

$$\forall (a,b,c) \in I^3,\ a < c < b \implies \frac{f(c)-f(a)}{c-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(c)}{b-c}\,.$$

Sur la figure ci-dessus, on a :

$$\text{pente}(AM) \leqslant \text{pente}(AB) \leqslant \text{pente}(MB)\,.$$

- **Croissance des pentes**
  Soit $f\colon I \to \mathbb{R}$. f est convexe sur I si et seulement si pour tout $a \in I$ la fonction
  $$x \mapsto \frac{f(x) - f(a)}{x - a}$$
  est croissante sur $I \setminus \{a\}$.
- On en déduit que si f est convexe sur I, elle est dérivable à droite et à gauche (donc continue) en tout point de $\overset{\circ}{I}$.
- **Fonctions convexes dérivables**
  - Si $f\colon I \to \mathbb{R}$ est dérivable sur I, elle est (strictement) convexe si et seulement si sa dérivée f' est (strictement) croissante sur I.
  - Si $f\colon I \to \mathbb{R}$ est dérivable et convexe, sa courbe représentative est au-dessus de ses tangentes, c'est-à-dire :
    $$\forall\, (a,x) \in I^2,\ f(x) \geqslant f(a) + (x - a)f'(a)\,.$$
  - Si $f\colon I \to \mathbb{R}$ est deux fois dérivable sur I, elle est convexe si et seulement si sa dérivée seconde f'' est à valeurs positives.

Thème 10 - Dérivation et intégration des fonctions vectorielles

On considère dans ce chapitre des applications définies sur un intervalle I de $\mathbb{R}$ (non réduit à un point), à valeurs dans un espace vectoriel normé E.
Les résultats énoncés généralisent ceux obtenus en première année pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Tous les espaces vectoriels normés intervenant dans ce chapitre seront supposés *de dimension finie*.

[S10.1] Dérivation d'une fonction vectorielle

- Soit f une application de I dans E, et $t_0 \in I$.
  On dit que f est dérivable en t_0 si

$$\lim_{\substack{t\to t_0\\ t\neq t_0}} \frac{1}{t-t_0}\big(f(t)-f(t_0)\big)$$

  existe dans E.
  Dans ce cas, cette limite se note $f'(t_0)$ ou $\dfrac{\mathrm{d}f}{\mathrm{d}t}(t_0)$ et s'appelle le vecteur dérivé de f en t_0.

  Comme pour les fonctions numériques, on définit de façon analogue les notions de vecteur dérivé à droite et de vecteur dérivé à gauche.

- **Développement limité**
  Dire que f est dérivable en t_0 peut aussi s'écrire :

$$\exists\, \ell \in E \text{ tel que } f(t) = f(t_0) + (t-t_0)\ell + o(t-t_0)$$

  où $o(t-t_0)$ désigne une fonction vectorielle de la forme $(t-t_0)\varepsilon(t)$ avec $\varepsilon : I \to E$ telle que $\lim\limits_{t\to t_0} \varepsilon(t) = 0$. Cette dernière expression s'appelle un développement limité de f au voisinage de t_0.
  Dans le cas où f est dérivable en t_0, ce développement limité peut aussi s'écrire sous la forme :

$$f(t_0+h) = f(t_0) + hf'(t_0) + o(h)\,.$$

- Si f est dérivable (respectivement dérivable à gauche, respectivement dérivable à droite) en t_0, alors f est continue (respectivement continue à gauche, respectivement continue à droite) en t_0.

  ✓ La réciproque de cette propriété est fausse. Il suffit par exemple de considérer l'application $f : x \mapsto |x|$ de $\mathbb{R}$ dans $\mathbb{R}$: f est évidemment continue sur $\mathbb{R}$ (elle est lipschitzienne de rapport 1), mais elle n'est pas dérivable en 0 (car $f'_g(0) = -1$ et $f'_d(0) = 1$).

- **Dérivation par coordonnées**
  On suppose E de dimension n, rapporté à une base $(e_1, \ldots, e_n)$.

Soit f une application de I dans E. Pour tout $t \in I$, on peut écrire :

$$f(t) = \sum_{i=1}^{n} f_i(t) e_i$$

où les $f_i : I \to \mathbb{K}$ sont les applications coordonnées de f.

Alors, f est dérivable en t_0 si et seulement si, pour tout $i \in [\![1\,;n]\!]$, f_i est dérivable en t_0, et on a alors :

$$f'(t_0) = \sum_{i=1}^{n} f_i'(t_0) e_i\,.$$

Exemples

- Une fonction $f : I \longrightarrow \mathbb{C}$ est dérivable en $t_0 \in I$ si et seulement si les fonctions $\mathcal{R}e\,(f)$ et $\mathcal{I}m\,(f)$ le sont, et on a alors $f'(t_0) = (\mathcal{R}e\,(f))'(t_0) + \mathrm{i}(\mathcal{I}m\,(f))'(t_0)$.
- Une application $A : \left\{ \begin{array}{ccc} I & \longrightarrow & \mathcal{M}_{p,q}(\mathbb{K}) \\ t & \longmapsto & (a_{i,j}(t))_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} \end{array} \right.$ est dérivable en t_0 si et seulement si pour tout $(i,j) \in [\![1\,;p]\!] \times [\![1\,;q]\!]$ les fonctions coefficients $t \mapsto a_{i,j}(t)$ le sont, et dans ce cas la matrice $A'(t_0)$ est la matrice dont les coefficients sont les $a'_{i,j}(t_0)$.

- **Dérivabilité sur un intervalle**
  - Une application $f : I \longrightarrow E$ est dite dérivable sur I si pour tout $t_0 \in I$, f est dérivable en t_0.

    Si tel est le cas, on peut définir l'application dérivée de f, notée f', qui à tout $t \in I$ associe le vecteur $f'(t) \in E$.

    L'ensemble des applications dérivables sur I à valeurs dans E sera noté $\mathscr{D}(I,E)$.
  - Une application $f : I \longrightarrow E$ est dite de classe $\mathscr{C}^1$ sur I si f est dérivable sur I et si f' est en plus continue sur I.

    On notera $\mathscr{C}^1(I,E)$ l'ensemble des applications de classe $\mathscr{C}^1$ sur I à valeurs dans E.

    ✓ Il existe des fonctions dérivables qui ne sont pas de classe $\mathscr{C}^1$! Autrement dit, l'inclusion $\mathscr{C}^1(I,E) \subset \mathscr{D}(I,E)$ est stricte.

[S10.2] Opérations sur les dérivées

- Soient f et g deux applications de I dans E, dérivables en t_0.
  Alors l'application $f+g$ est dérivable en t_0, et $(f+g)'(t_0) = f'(t_0) + g'(t_0)$.
- Soit f une application de I dans E, et λ une application de I dans $\mathbb{K}$, dérivables en t_0.
  Alors l'application $\lambda.f$ est dérivable en t_0 et $(\lambda.f)'(t_0) = \lambda'(t_0)f(t_0) + \lambda(t_0)f'(t_0)$.
  En particulier, si $\alpha \in \mathbb{K}$ est constant, αf est dérivable en t_0 et $(\alpha f)'(t_0) = \alpha f'(t_0)$.
  Il en résulte que l'ensemble $\mathscr{D}(I,E)$ est un sous-espace vectoriel de $\mathscr{C}^0(I,E)$, et l'application $f \mapsto f'$ est linéaire de $\mathscr{D}(I,E)$ dans $\mathcal{A}(I,E)$.

- Soient E, F deux espaces vectoriels normés, f une application d'un intervalle I de $\mathbb{R}$ à valeurs dans E, et u une application linéaire de E dans F.
  Si f est dérivable en t_0, $u \circ f$ l'est aussi et : $(u \circ f)'(t_0) = u[f'(t_0)]$.

- Soient E, F, G trois espaces vectoriels normés, et $B\colon E \times F \longrightarrow G$ une application bilinéaire.
  Soient $f\colon I \longrightarrow E$ et $g\colon I \longrightarrow F$ deux applications dérivables en $t_0 \in I$.
  Alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & G \\ t & \longmapsto & B\big(f(t),g(t)\big) \end{cases}$ est dérivable en t_0 et

$$\varphi'(t_0) = B\big(f'(t_0),g(t_0)\big) + B\big(f(t_0),g'(t_0)\big).$$

  Exemple : soit E un espace préhilbertien réel, muni d'un produit scalaire noté $\langle \cdot \,|\, \cdot \rangle$.
  Si f et g sont deux applications définies sur I, à valeurs dans E et dérivables, alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \langle f(t) \,|\, g(t) \rangle \end{cases}$ est dérivable et :

$$\langle f \,|\, g \rangle' = \langle f' \,|\, g \rangle + \langle f \,|\, g' \rangle .$$

  On en déduit ensuite que, si $f\colon I \longrightarrow E$ est dérivable en t_0 et si $f(t_0) \neq 0$, l'application $t \mapsto \|f(t)\|$ est dérivable en t_0, et :

$$\|f\|'(t_0) = \frac{\langle f(t_0) \,|\, f'(t_0) \rangle}{\|f(t_0)\|} .$$

- Le théorème précédent peut se généraliser.
  Soient $E_1, \ldots, E_n$ et F des espaces vectoriels normés, et $g\colon E_1 \times E_2 \times \cdots \times E_n \longrightarrow F$ une application n-linéaire.
  Pour tout $i \in [\![1\,;n]\!]$, soit $f_i\colon I \longrightarrow E_i$ une application dérivable sur I.
  Alors l'application $\varphi\colon \begin{cases} I & \longrightarrow & F \\ t & \longmapsto & g\big(f_1(t),f_2(t),\ldots,f_n(t)\big) \end{cases}$ est dérivable sur I et, pour tout $t \in I$:

$$\varphi'(t) = \sum_{i=1}^{n} g\big(f_1(t),\ldots,f_{i-1}(t),f_i'(t),f_{i+1}(t),\ldots,f_n(t)\big).$$

  Exemples

  - Si $f_1, \ldots, f_n$ sont n applications dérivables sur I et à valeurs dans une algèbre normée, leur produit $g\colon t \mapsto f_1(t)f_2(t)\cdots f_n(t)$ est dérivable sur I et pour tout $t \in I$:

$$g'(t) = f_1'(t)f_2(t)\cdots f_n(t) + f_1(t)f_2'(t)f_3(t)\cdots f_n(t) + \cdots + f_1(t)f_2(t)\cdots f_{n-1}(t)f_n'(t) .$$

  - On sait que l'application $\det\colon \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow & K \\ M & \longmapsto & \det M \end{cases}$ est une application n-linéaire des colonnes de la matrice M.

Soit alors $M : \begin{cases} I & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto & M(t) \end{cases}$ une application dérivable sur I.

Si l'on note $C_1(t), \ldots, C_n(t)$ les colonnes de la matrice $M(t)$, alors les applications $t \mapsto C_i(t)$ sont dérivables sur I (car leurs fonctions coordonnées le sont). Par suite, l'application $\varphi : t \mapsto \det(M(t))$ est dérivable sur I et :

$$\varphi'(t) = \sum_{j=1}^{n} \det\big(C_1(t), \ldots, C_{j-1}(t), C_j'(t), C_{j+1}(t), \ldots, C_n(t)\big).$$

- **Fonction composée**

  Soit $\varphi : I \longrightarrow J$ et $f : J \longrightarrow E$, où I et J sont des intervalles de $\mathbb{R}$ et E un espace vectoriel normé.

  Soit $t_0 \in I$. Si φ est dérivable en t_0 et si f est dérivable en $\varphi(t_0)$, alors $f \circ \varphi$ est dérivable en t_0 et :

  $$(f \circ \varphi)'(t_0) = \varphi'(t_0) f'[\varphi(t_0)] .$$

[S10.3] Dérivées successives

- Soit f une application de I dans E. On peut définir par récurrence, si elles existent, les dérivées successives de f de la façon suivante :
  - on pose $f^{(0)} = f$ (dérivée d'ordre zéro) ;
  - pour $k \in \mathbb{N}^*$, on dira que f est k fois dérivable sur I si $f^{(k-1)}$ est dérivable sur I et on pose alors $f^{(k)} = \big(f^{(k-1)}\big)'$ (dérivée d'ordre k).

  On notera $\mathscr{D}^k(I,E)$ l'ensemble des applications k fois dérivables sur I à valeurs dans E.

  Pour $k \in \mathbb{N}^*$, on dira que f est de classe $\mathscr{C}^k$ sur I si f est k fois dérivable sur I et si $f^{(k)}$ est continue sur I ; on note $\mathscr{C}^k(I,E)$ l'ensemble des applications de classe $\mathscr{C}^k$ sur I à valeurs dans E.

  Enfin on dira que f est de classe $\mathscr{C}^\infty$ sur I si elle est de classe $\mathscr{C}^k$ pour tout $k \in \mathbb{N}$, et on notera $\mathscr{C}^\infty(I,E)$ l'ensemble des applications de classe $\mathscr{C}^\infty$ sur I à valeurs dans E.

- $\mathscr{D}^k(I,E)$, $\mathscr{C}^k(I,E)$ et $\mathscr{C}^\infty(I,E)$ sont des sous-espaces vectoriels de $\mathcal{A}(I,E)$.

  Si $f, g \in \mathscr{D}^k(I,E)$ et si $\lambda, \mu \in \mathbb{K}$, on a : $(\lambda f + \mu g)^{(k)} = \lambda f^{(k)} + \mu g^{(k)}$.

- On suppose E de dimension n, rapporté à une base $(e_1, \ldots, e_n)$.

  Soit f une application de I dans E. Pour tout $t \in I$, on peut écrire $f(t) = \sum_{i=1}^{n} f_i(t) e_i$ (les $f_i : I \to \mathbb{K}$ sont les applications coordonnées).

  Alors, f est k-fois dérivable (respectivement de classe $\mathscr{C}^k$) sur I si et seulement si, pour tout $i \in [\![1\,;n]\!]$, f_i est k fois dérivable (respectivement de classe $\mathscr{C}^k$) sur I, et on a alors :

  $$\forall\, t \in I,\ f^{(k)}(t) = \sum_{i=1}^{n} {f_i}^{(k)}(t) e_i .$$

- **Formule de Leibniz**

  Soient E, F, G trois espaces vectoriels normés, et $B\colon E\times F\longrightarrow G$ une application bilinéaire.

  Soient $f\colon I\longrightarrow E$ et $g\colon I\longrightarrow F$ deux applications n fois dérivables (respectivement de classe $\mathscr{C}^n$) sur I .

  Alors l'application $\varphi\colon\begin{cases} I & \longrightarrow & G\\ t & \longmapsto & B\big(f(t),g(t)\big)\end{cases}$ est n fois dérivable (respectivement de classe $\mathscr{C}^n$) sur I, et l'on a :

$$\forall\, t\in I,\ \varphi^{(n)}(t)=\sum_{k=0}^{n}\binom{n}{k}B\big(f^{(k)}(t),g^{(n-k)}(t)\big)\,.$$

- **Conséquences**
  - $\mathscr{D}^n(I,\mathbb{K})$ et $\mathscr{C}^n(I,\mathbb{K})$ sont des sous-algèbres de $\mathcal{A}(I,\mathbb{K})$.
  - Soit $\varphi\colon I\longrightarrow J$ et $f\colon J\longrightarrow E$, où I et J sont des intervalles de $\mathbb{R}$ et E un espace vectoriel normé.

    Si $\varphi\in\mathscr{C}^n(I,J)$ et $f\in\mathscr{C}^n(J,E)$, alors $f\circ\varphi\in\mathscr{C}^n(I,E)$.

[S10.4] Applications de classe $\mathscr{C}^k$ par morceaux

- **Définitions**
  - Soit $[a\,;b]$ $(a<b)$ un segment de $\mathbb{R}$, $k\in\mathbb{N}\cup\{\infty\}$ et f une application de $[a\,;b]$ dans un espace vectoriel normé E.

    On dira que f est de classe $\mathscr{C}^k$ par morceaux sur $[a\,;b]$ s'il existe une subdivision $a=x_0<x_1<\ldots<x_{n-1}<x_n=b$ telle que la restriction de f à chaque intervalle $]x_{i-1}\,;x_i[$ pour $1\leqslant i\leqslant n$ se prolonge en une application de classe $\mathscr{C}^k$ sur $[x_{i-1}\,;x_i]$. Une telle subdivision est dite adaptée à f.
  - Soit I un intervalle quelconque de $\mathbb{R}$, et f une application de I dans un espace vectoriel normé E.

    On dira que f est de classe $\mathscr{C}^k$ par morceaux sur I si sa restriction à tout segment $[a\,;b]$ inclus dans I est de classe $\mathscr{C}^k$ par morceaux.
  - On notera $\mathscr{CM}^k(I,E)$ l'ensemble des applications de classe $\mathscr{C}^k$ par morceaux de I dans E ; dans le cas $k=0$, on parle simplement d'application continue par morceaux et on notera $\mathscr{CM}(I,E)$ l'ensemble des applications continues par morceaux de I dans E.
- **Exemples**
  - La fonction « partie entière » est continue par morceaux sur $\mathbb{R}$.
  - La fonction tan n'est pas continue par morceaux sur $\mathbb{R}$.
  - La fonction $x\mapsto\sqrt{x}$ est $\mathscr{C}^1$ par morceaux sur $\mathbb{R}_+^*$ mais pas sur $\mathbb{R}_+$.
- L'ensemble $\mathscr{CM}^k(I,E)$ des applications de classe $\mathscr{C}^k$ par morceaux de I dans E est un sous-espace vectoriel de $\mathcal{A}(I,E)$.

[S10.5] Arcs paramétrés

Nous nous contentons ici de rappeler les définitions de base. Les méthodes pratiques d'étude d'un arc paramétré seront détaillées dans le chapitre Savoir-Faire page 430.

Pour toute la suite, on se donne un entier $n \geqslant 2$; l'ensemble $\mathbb{R}^n$ peut être considéré comme un espace vectoriel (un n-uplet $(x_1,\ldots,x_n)$ étant alors un vecteur $\overrightarrow{V}$) ou comme un espace affine (une origine O étant choisie, le n-uplet $(x_1,\ldots,x_n)$ représente les coordonnées d'un point M tel que $\overrightarrow{OM} = \overrightarrow{V}$).

- On appelle arc paramétré (ou courbe paramétrée) de $\mathbb{R}^n$ un couple (I,γ) où I est un intervalle de $\mathbb{R}$ non réduit à un point et γ une application de I dans $\mathbb{R}^n$. L'arc est dit de classe $\mathscr{C}^k$ si γ est de classe $\mathscr{C}^k$.

  Pour tout $t \in I$ on notera $M(t)$ le point de $\mathbb{R}^n$ tel que $\overrightarrow{OM(t)} = \gamma(t)$.
  L'ensemble :
  $$\Gamma = \{M(t) \in \mathbb{R}^n \mid t \in I\}$$
  des points de $\mathbb{R}^n$ atteints est appelé support de l'arc.

  En interprétant le paramètre t comme le temps, un arc paramétré représente le mouvement d'un mobile $M(t)$; le support de l'arc est la trajectoire du mobile. Si γ est de classe $\mathscr{C}^k$ avec $k \geqslant 2$,

  $\gamma'(t) = \dfrac{\mathrm{d}\overrightarrow{OM}}{\mathrm{d}t}$ est le vecteur vitesse et $\gamma''(t) = \dfrac{\mathrm{d}^2\overrightarrow{OM}}{\mathrm{d}t^2}$ le vecteur accélération.

- **Tangente en un point d'un arc paramétré**

  Soit (I,γ) un arc paramétré de classe $\mathscr{C}^k$, et $t_0 \in I$. On suppose qu'il existe un entier $i \in [\![1\,;k]\!]$ tel que $\gamma^{(i)}(t_0) \neq 0$ et on note p le plus petit de ces entiers.
  On a alors le développement limité :
  $$\gamma(t_0+h) = \gamma(t_0) + h^p\,\frac{\gamma^{(p)}(t_0)}{p!} + o(h^p)$$
  de sorte que : $\displaystyle\lim_{h\to 0} \frac{\overrightarrow{M(t_0)M(t_0+h)}}{h^p} = \frac{\gamma^{(p)}(t_0)}{p!}$.

  On appelle donc tangente à l'arc au point $M(t_0)$ de paramètre t_0 la droite passant par $M(t_0)$ et dirigée par le vecteur $\gamma^{(p)}(t_0)$ c'est-à-dire l'ensemble :
  $$\left\{M(t_0) + \lambda\gamma^{(p)}(t_0) \;\middle|\; \lambda \in \mathbb{R}\right\}.$$
  Le point $M(t_0)$ de paramètre t_0 est dit régulier si $p = 1$, c'est-à-dire si $\gamma'(t_0) \neq 0$. Il est dit stationnaire sinon, c'est-à-dire si $\gamma'(t_0) = 0$.
  Un arc est dit régulier si tous ses points sont réguliers.

- **Longueur d'un arc paramétré**

  On suppose ici $\mathbb{R}^n$ muni de sa structure euclidienne canonique.
  Soit (I,γ) un arc paramétré de classe $\mathscr{C}^1$ à valeurs dans $\mathbb{R}^n$, et $t_0 < t_1$ dans I.
  La longueur de l'arc entre les points $M(t_0)$ et $M(t_1)$ de paramètres t_0 et t_1 est le réel positif :
  $$\overbrace{M(t_0)M(t_1)} = \int_{t_0}^{t_1} \|\gamma'(t)\|\,\mathrm{d}t\,.$$

En particulier, si l'arc est à valeurs dans $\mathbb{R}^2$ on obtient la formule :

$$\overset{\frown}{M(t_0)M(t_1)} = \int_{t_0}^{t_1} \sqrt{x'^2(t) + y'^2(t)}\,\mathrm{d}t\,.$$

[S10.6] Intégrale d'une fonction continue par morceaux sur un segment

Les applications considérées par la suite sont définies sur un *segment* $[a\,;b]$ de $\mathbb{R}$ $(a < b)$ et à valeurs dans un espace vectoriel normé de dimension finie E.

- **Fonctions en escalier**
  - Une application $f\colon [a\,;b] \to E$ où E est un $\mathbb{K}$-espace vectoriel normé est dite en escalier sur $[a\,;b]$ s'il existe une subdivision $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ de l'intervalle $[a\,;b]$ telle que la restriction de f à chaque intervalle $]x_{i-1}\,;x_i[$ pour $1 \leqslant i \leqslant n$ soit constante. Une telle subdivision est dite adaptée à f.
  - Si I est un intervalle quelconque de $\mathbb{R}$, f est dite en escalier sur I si elle l'est sur tout segment inclus dans I.
  - L'ensemble $\mathscr{E}sc(I,E)$ des fonctions en escalier de I dans E est un sous-espace vectoriel de $\mathcal{A}(I,E)$.
- **Intégrale d'une fonction en escalier sur un segment**

  Soit $f \in \mathscr{E}sc([a\,;b],E)$ et $\sigma = (x_0,\ldots,x_n)$ une subdivision adaptée à f, c'est-à-dire :

  $$\forall\, i \in [\![1\,;n]\!],\quad \forall\, t \in\,]x_{i-1}\,;x_i[,\quad f(t) = c_i$$

  où les c_i sont des éléments de E. On pose :

  $$I_\sigma(f) = \sum_{i=1}^{n}(x_i - x_{i-1})c_i\,.$$

  Alors le vecteur $I_{\sigma(f)}$ est indépendant de la subdivision σ *adaptée* à f choisie.

  On l'appelle intégrale de f sur $[a\,;b]$. Il est noté $\displaystyle\int_{[a;b]} f$ ou $\displaystyle\int_a^b f$ ou $\displaystyle\int_a^b f(t)\,\mathrm{d}t$.
- **Intégrale d'une fonction continue par morceaux**

  On sait qu'une fonction continue par morceaux sur un segment y est limite uniforme d'une suite de fonctions en escalier (voir **[S12.6]**).

  Soit $f \in \mathscr{CM}([a\,;b],E)$. Pour toute suite $(\varphi_n)_{n\in\mathbb{N}}$ de fonctions en escalier qui converge uniformément vers f sur $[a\,;b]$, la suite $\left(\displaystyle\int_a^b \varphi_n\right)_{n\in\mathbb{N}}$ converge dans E.

  De plus, sa limite ne dépend que de f, et non de la suite $(\varphi_n)_{n\in\mathbb{N}}$ choisie. Cette limite s'appelle l'intégrale de f sur $[a\,;b]$, et est notée $\displaystyle\int_a^b f$ ou $\displaystyle\int_{[a;b]} f$.

[S10.7] Propriétés de l'intégrale

- **Linéarité**

  L'application $I : \begin{cases} \mathscr{CM}([a\,;b],E) & \longrightarrow & E \\ f & \longmapsto & \displaystyle\int_{[a;b]} f \end{cases}$ est linéaire.

- Soient E, F deux espaces vectoriels normés.
  Soit $f \in \mathscr{CM}([a\,;b],E)$ et u une application linéaire de E dans F.
  Alors, $u \circ f$ est continue par morceaux de $[a,b]$ dans F et l'on a :

$$\int_{[a,b]} u \circ f = u\left(\int_{[a,b]} f\right).$$

- **Invariance**

  Soient f, $g \in \mathscr{CM}([a\,;b],E)$. Si f et g ne diffèrent qu'en un nombre fini de points, alors $\displaystyle\int_a^b f = \int_a^b g$.

- Soit $f \in \mathscr{CM}([a\,;b],E)$. Alors la fonction $\|f\| : t \mapsto \|f(t)\|$ est continue par morceaux sur $[a\,;b]$ et :

$$\left\|\int_a^b f\right\| \leqslant \int_a^b \|f\| \leqslant (b-a)\,\|f\|_\infty\,,$$

  où $\|f\|_\infty$ désigne comme d'habitude le réel $\sup\{\|f(t)\|_E\,,\, t \in [a\,;b]\}$.

- **Relation de Chasles**

  Soit $f \in \mathscr{CM}([a\,;b],E)$ et $c \in]a\,;b[$.
  Alors $f|_{[a;c]}$ et $f|_{[c;b]}$ sont continues par morceaux et l'on a :

$$\int_a^b f = \int_a^c f + \int_c^b f\,.$$

- **Utilisation des fonctions coordonnées dans une base**

  Soit $\mathscr{B} = (e_1,\ldots,e_n)$ une base de E. Si $f \in \mathscr{CM}([a\,;b],E)$ on a, pour tout $t \in [a\,;b]$, $f(t) = \displaystyle\sum_{i=1}^{n} f_i(t)e_i$, où les f_i sont des applications de $[a\,;b]$ dans $\mathbb{K}$.
  Alors les f_i sont continues par morceaux sur $[a\,;b]$ et l'on a :

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{i=1}^{n}\left(\int_a^b f_i(t)\,\mathrm{d}t\right)e_i\,.$$

- **Cas des fonctions à valeurs réelles**

  . Positivité

  – Si f est une application continue par morceaux sur $[a\,;b]$, *à valeurs réelles positives*, alors $\displaystyle\int_a^b f \geqslant 0$.

  – Si f, $g \in \mathscr{CM}([a\,;b],\mathbb{R})$ sont telles que $f \leqslant g$ sur $[a\,;b]$, alors $\displaystyle\int_a^b f \leqslant \int_a^b g$.

– Si f est une fonction *continue* sur $[a\,;b]$, avec $a < b$, *à valeurs réelles positives*, telles que $\int_a^b f = 0$, alors $f = 0$ sur $[a\,;b]$.

. **Inégalité de Cauchy-Schwarz**
Pour $f,\, g \in \mathscr{CM}([a\,;b],\mathbb{R})$ on a :

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leqslant \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\sqrt{\int_a^b g(t)^2\,\mathrm{d}t}\,.$$

De plus, lorsque f et g sont *continues*, il y a égalité si et seulement si la famille (f,g) est liée.

[S10.8] Sommes de Riemann

- **Définition**
Soit $\sigma = (x_0,\ldots,x_n)$ une subdivision de $[a\,;b]$.
On appelle pas de cette subdivision le réel :

$$\max\{x_{i+1} - x_i,\, i \in [\![0\,;n-1]\!]\}\,.$$

On appelle famille subordonnée à σ toute famille $c = (c_i)_{1\leqslant i\leqslant n}$ telle que $c_i \in [x_{i-1}\,;x_i]$ pour tout $i \in [\![1\,;n]\!]$.
Si f est une fonction continue par morceaux sur $[a\,;b]$, à valeurs dans E, le vecteur $S(f,\sigma,c) = \sum_{i=1}^{n}(x_i - x_{i-1})f(c_i)$ s'appelle une somme de Riemann relative à f, σ, c.

- **Théorème**
Soit f une fonction continue par morceaux sur $[a\,;b]$ à valeurs dans E.
Alors, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour toute subdivision σ de $[a\,;b]$ de pas inférieur à α et toute suite subordonnée c, on ait :

$$\left\|S(f,\sigma,c) - \int_a^b f\right\| < \epsilon$$

(autrement dit, les sommes de Riemann associées à f tendent vers $\int_a^b f$ lorsque le pas de la subdivision tend vers 0).

- **Cas particulier**
On utilise le plus souvent le théorème précédent dans le cas où σ est une subdivision régulière de pas $\frac{b-a}{n}$, c'est-à-dire $x_i = a + i\frac{b-a}{n}\cdot$ On obtient dans ce cas :

$$S(f,\sigma,c) = \frac{b-a}{n}\sum_{i=1}^{n} f(c_i) \quad (c_i \in [x_{i-1}\,;x_i])\,.$$

En choisissant $c_i = x_i$ ($c_i = x_{i-1}$ ou $c_i = \frac{x_i + x_{i-1}}{2}$ sont aussi des choix possibles), on obtient :

$$\lim_{n\to+\infty}\frac{b-a}{n}\sum_{i=1}^{n} f\left(a + i\frac{b-a}{n}\right) = \int_a^b f\,.$$

[S10.9] Primitives

- **Définitions**
  - Soit f une application *continue* sur un intervalle $I \subset \mathbb{R}$, à valeurs dans E.
    On appelle primitive de f toute application $F\colon I \to E$, de classe $\mathscr{C}^1$, telle que $F' = f$.
  - Soit f une application *continue par morceaux* sur un intervalle $I \subset \mathbb{R}$, à valeurs dans E.
    On appelle alors primitive de f une application $F\colon I \to E$ telle que :
    - F est continue et de classe $\mathscr{C}^1$ par morceaux sur I ;
    - et $F'(t) = f(t)$ en tout point t de I où f est continue.
- Si $f \in \mathscr{CM}(I,E)$, deux primitives quelconques de f diffèrent d'une constante.
- **Théorème fondamental**
  Si $f \in \mathscr{CM}(I,E)$, et si $a \in I$, l'application $F_a\colon \left\{ \begin{array}{ccc} I & \longrightarrow & E \\ x & \longmapsto & \displaystyle\int_a^x f(t)\,\mathrm{d}t \end{array} \right.$ est une primitive de f.
  C'est l'unique primitive de f qui s'annule en a.
- **Conséquences**
  - Si f est continue par morceaux sur $[a\,;b]$ et si F est une primitive de f, on a :
    $$\int_a^b f = F(b) - F(a) \quad \text{ce que l'on écrit :} \quad \int_a^b f = \big[F(t)\big]_a^b .$$
  - Si f est *continue* et *de classe $\mathscr{C}^1$ par morceaux* sur $[a\,;b]$, on a :
    $$\int_a^b f' = f(b) - f(a) .$$
    (avec, ici, un abus de notation, f' n'étant définie que sur $[a\,;b]$ privé d'un nombre fini de points.)
  - **Inégalité des accroissements finis**
    Si f est continue sur $[a\,;b]$ et de classe $\mathscr{C}^1$ sur $]a\,;b[$ et s'il existe un réel M tel que : $\forall\, t \in\,]a\,;b[,\ \|f'(t)\| \leqslant M$, on a :
    $$\|f(b) - f(a)\| \leqslant M(b-a) .$$
- **Généralisation : intégrale fonction de ses bornes**
  Soit f une application continue de I dans E, et u, v deux applications de classe $\mathscr{C}^1$ définies sur un intervalle J, à valeurs dans I.
  Alors, la fonction $F\colon \left\{ \begin{array}{ccc} J & \longrightarrow & E \\ x & \longmapsto & \displaystyle\int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t \end{array} \right.$ est de classe $\mathscr{C}^1$ sur J et :
  $$\forall\, x \in J,\ F'(x) = v'(x).f\big[v(x)\big] - u'(x).f\big[u(x)\big] .$$

[S10.10] Intégration par parties

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels normés, et B une application bilinéaire de $E \times F$ dans G.
Soient $f\colon I \to E$ et $g\colon I \to F$ des applications continues et de classe $\mathscr{C}^1$ par morceaux sur un intervalle I de $\mathbb{R}$. Alors :

$$\forall\,(a,b) \in I^2,\ \int_a^b B\big(f'(t),g(t)\big)\,\mathrm{d}t = \Big[B\big(f(t),g(t)\big)\Big]_a^b - \int_a^b B\big(f(t),g'(t)\big)\,\mathrm{d}t\,.$$

[S10.11] Changement de variable

- *Cas d'une fonction continue*
  Soit $f \in \mathscr{C}^0(I,E)$ et $\varphi \in \mathscr{C}^1([a\,;b],\mathbb{R})$ telle que $\varphi([a\,;b]) \subset I$. Alors :

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,\mathrm{d}t = \int_a^b \varphi'(u).f[\varphi(u)]\,\mathrm{d}u\,.$$

- *Extension aux fonctions continues par morceaux*
  Soit $f \in \mathscr{CM}(I,E)$ et $\varphi \in \mathscr{C}^1([a\,;b],\mathbb{R})$, strictement monotone, telle que $\varphi([a\,;b]) \subset I$. Alors :

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,\mathrm{d}t = \int_a^b \varphi'(u).f[\varphi(u)]\,\mathrm{d}u\,.$$

[S10.12] Formules de Taylor

- **Formule de Taylor avec reste intégrale**
  Soit $f\colon I \longrightarrow E$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I et de classe $\mathscr{C}^{n+1}$ par morceaux sur I. Alors, pour tous a, $b \in I$:

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t\,.$$

- **Inégalité de Taylor-Lagrange**
  Soit $f\colon I \longrightarrow E$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I et de classe $\mathscr{C}^{n+1}$ par morceaux sur I. Alors, pour tous a, $b \in I$:

$$\left\| f(b) - \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) \right\| \leqslant \frac{|b-a|^{n+1}}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty^{[a;b]}\,.$$

- **Formule de Taylor-Young**
  Soit $f\colon I \longrightarrow E$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N})$ sur I. Alors, pour tous a, $x \in I$:

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + (x-a)^n \varepsilon(x-a)\,,$$

  où ε est une fonction à valeurs dans E telle que $\lim\limits_{h \to 0} \varepsilon(h) = 0$.

[S10.13] Primitives usuelles

Dans le tableau ci-dessous, on rappelle quelques primitives de fonctions usuelles. Les domaines de définition ne sont pas mentionnés. Le réel a qui intervient est supposé non nul.

Fonction	Primitive	Fonction	Primitive
$(ax+b)^\alpha$ $(\alpha \neq -1)$	$\dfrac{(ax+b)^{\alpha+1}}{a(\alpha+1)}$	$\operatorname{sh}(ax+b)$	$\dfrac{1}{a}\operatorname{ch}(ax+b)$
$\dfrac{1}{ax+b}$	$\dfrac{1}{a}\ln\lvert ax+b\rvert$	$\operatorname{ch}(ax+b)$	$\dfrac{1}{a}\operatorname{sh}(ax+b)$
e^{ax+b}	$\dfrac{1}{a}\mathrm{e}^{ax+b}$	$\dfrac{1}{\operatorname{sh} x}$	$\ln\left\lvert\operatorname{th}\dfrac{x}{2}\right\rvert$
$\ln x$	$x\ln x - x$	$\dfrac{1}{\operatorname{ch} x}$	$2\operatorname{Arctan}(\mathrm{e}^x)$
$\sin(ax+b)$	$-\dfrac{1}{a}\cos(ax+b)$	$\dfrac{1}{\operatorname{sh}^2 x}$	$\coth x$
$\cos(ax+b)$	$\dfrac{1}{a}\sin(ax+b)$	$\dfrac{1}{\operatorname{ch}^2 x}$	$\operatorname{th} x$
$\dfrac{1}{\sin x}$	$\ln\left\lvert\tan\dfrac{x}{2}\right\rvert$	$\operatorname{th} x$	$\ln(\operatorname{ch} x)$
$\dfrac{1}{\cos x}$	$\ln\left\lvert\tan\left(\dfrac{x}{2}+\dfrac{\pi}{4}\right)\right\rvert$	$\dfrac{1}{x^2+a^2}$	$\dfrac{1}{a}\operatorname{Arctan}\left(\dfrac{x}{a}\right)$
$\dfrac{1}{\sin^2 x}$	$-\cot x$	$\dfrac{1}{x^2-a^2}$	$\dfrac{1}{2a}\ln\left\lvert\dfrac{x-a}{x+a}\right\rvert$
$\dfrac{1}{\cos^2 x}$	$\tan x$	$\dfrac{1}{\sqrt{x^2+a^2}}$	$\ln(x+\sqrt{x^2+a^2})$
$\tan x$	$-\ln\lvert\cos x\rvert$	$\dfrac{1}{\sqrt{x^2-a^2}}$	$\ln\lvert x+\sqrt{x^2-a^2}\rvert$
$\tan^2 x$	$\tan x - x$	$\dfrac{1}{\sqrt{a^2-x^2}}$	$\operatorname{Arcsin}\left(\dfrac{x}{\lvert a\rvert}\right)$

On retiendra qu'une primitive de $\dfrac{1}{x^2-a^2}$ (qui intervient fréquemment) se retrouve facilement en écrivant $\dfrac{1}{x^2-a^2} = \dfrac{1}{2a}\left(\dfrac{1}{x-a} - \dfrac{1}{x+a}\right)$.

Thème 11 - Séries et familles sommables

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et $(E, \|\cdot\|)$ est un $\mathbb{K}$-espace vectoriel normé.

[S11.1] Définition de la convergence d'une série

- Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. On appelle série de terme général u_n et l'on note $\sum_{n\in\mathbb{N}} u_n$ la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ définie par :

$$\forall\, n \in \mathbb{N},\ S_n = \sum_{k=0}^{n} u_k\,.$$

On dit alors que la série $\sum u_n$ converge si la suite $(S_n)_{n\in\mathbb{N}}$ converge dans E.

Dans ce cas, on appelle somme de la série $\sum u_n$ la limite de la suite $(S_n)_{n\in\mathbb{N}}$; cette somme est notée $\sum_{n=0}^{+\infty} u_n$.

✓ La nature (convergence ou divergence) de la série n'est pas modifiée si l'on part d'un entier $n_0 \neq 0$ (mais, en cas de convergence, la valeur de la somme peut être modifiée).

✓ On fera particulièrement attention aux notations : l'écriture $\sum u_n$ ou $\sum_{n\in\mathbb{N}} u_n$ est une simple abréviation pour dire « la série de terme général u_n », et ne doit pas être utilisée dans des calculs ; l'écriture $\sum_{n=0}^{N} u_n$ désigne une « vraie » somme, et la notation $\sum_{n=0}^{+\infty} u_n$, *qui ne doit être utilisée que si la série converge*, désigne la limite de la suite des sommes partielles.

[S11.2] Séries télescopiques

Il s'agit de séries de la forme $\sum v_n$, où $v_n = u_n - u_{n-1}$ $(n \geqslant 1)$.

Les sommes partielles sont alors : $V_n = \sum_{k=1}^{n} v_k = u_n - u_0$ donc on peut énoncer le théorème suivant :

La *série* $\sum_{n\in\mathbb{N}^*} (u_n - u_{n-1})$ converge si et seulement si la *suite* (u_n) converge et, dans ce cas : $\sum_{k=1}^{+\infty} (u_k - u_{k-1}) = \lim_{n\to+\infty} u_n - u_0$.

✓ Cette propriété, qui peut sembler simpliste, est en réalité très importante : elle permet non seulement de calculer la somme de certaines séries, mais aussi, dans certains cas, d'étudier la convergence d'une suite en la ramenant à celle d'une série, donc en utilisant les nombreux outils supplémentaires disponibles sur les séries.

[S11.3] Condition nécessaire de convergence d'une série

Si la série $\sum u_n$ converge alors $\lim\limits_{n\to+\infty} u_n = 0_E$.

✓ Cela équivaut au résultat essentiel suivant : une série dont le terme général ne tend pas vers 0 est divergente ; on parle de divergence grossière.

Mais attention : on ne peut rien dire a priori d'une série dont le terme général tend vers 0. Il faut d'ailleurs retenir l'exemple de la série harmonique $\sum\limits_{n\geqslant 1} \frac{1}{n}$, dont le terme général tend vers zéro mais qui est divergente.

[S11.4] Opérations algébriques sur les séries

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites d'éléments de E et λ un scalaire.

- Si les séries $\sum u_n$ et $\sum v_n$ convergent alors la série $\sum(\lambda u_n + v_n)$ converge et :

$$\sum_{n=0}^{+\infty}(\lambda u_n + v_n) = \lambda\sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n\,.$$

- Si la série $\sum u_n$ converge et si la série $\sum v_n$ diverge, alors la série $\sum(u_n + v_n)$ diverge.

  ✓ On ne peut rien dire en général de la somme de deux séries divergentes.

- On suppose ici E de dimension finie $p \in \mathbb{N}^*$, muni d'une base $\mathscr{B} = (e_1, \ldots, e_p)$. Si $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, on peut considérer ses suites coordonnées $(u_n^{(i)})_{n\in\mathbb{N}}$ définies par :

$$\forall\, n \in \mathbb{N},\ u_n = \sum_{i=1}^{p} u_n^{(i)} e_i\,.$$

  Alors la série $\sum u_n$ converge dans E si et seulement si pour tout $i \in [\![1\,;n]\!]$ la série numérique $\sum u_n^{(i)}$ converge dans $\mathbb{K}$.
  Et dans ce cas l'on a :

$$\sum_{n=0}^{+\infty} u_n = \sum_{i=1}^{p}\left(\sum_{n=0}^{+\infty} u_n^{(i)}\right) e_i\,.$$

  ✓ Ce résultat permet de ramener l'étude d'une série vectorielle, dans le cas de la dimension finie, à celle de ses séries coordonnées. Cependant, pour étudier une série $\sum u_n$ à valeurs dans E, on préfèrera dans la plupart des cas démontrer l'absolue convergence (voir ci-après) en considérant la série des normes $\sum \|u_n\|$.

[S11.5] Convergence absolue

Une série $\sum u_n$ de vecteurs de l'espace vectoriel normé $(E, \|\cdot\|)$ est dite absolument convergente si la série de nombres réels positifs $\sum \|u_n\|$ est convergente.

Théorème : si E est de dimension finie, toute série $\sum u_n$ absolument convergente est convergente.
Dans ce cas on a de plus :

$$\left\| \sum_{n=0}^{+\infty} u_n \right\| \leqslant \sum_{n=0}^{+\infty} \|u_n\| \ .$$

✓ La réciproque de ce théorème est fausse ; il faut connaître le contre-exemple classique : la série harmonique alternée $\sum_{n\geqslant 1} \frac{(-1)^n}{n}$ est convergente alors que la série harmonique $\sum_{n\geqslant 1} \frac{1}{n}$ est divergente.

[S11.6] Règles de comparaison pour les séries à termes réels positifs

Nous rappelons ici les règles de comparaison vues en 1re année pour les séries à termes réels positifs.
Pour une série $\sum u_n$ à valeurs dans un espace vectoriel normé E, on pourra chercher à utiliser ces règles en les appliquant à la série $\sum \|u_n\|$ afin de démontrer l'absolue convergence de la série (donc la convergence lorsque E est de dimension finie).

- **Règle de comparaison n° 1**
  Une série à termes réels positifs est convergente si et seulement si la suite de ses sommes partielles est majorée.
- **Règle de comparaison n° 2**
  Soit (u_n) et (v_n) deux suites de nombres réels telles que :

  $$0 \leqslant u_n \leqslant v_n \quad \text{à partir d'un certain rang.}$$

  – Si la série $\sum v_n$ converge, la série $\sum u_n$ converge.
  – Si la série $\sum u_n$ diverge, la série $\sum v_n$ diverge.
- **Règle de comparaison n° 3**
  Soit (u_n) une suite à valeurs réelles et (v_n) une suite de nombres réels *positifs*.
  On suppose qu'il existe un réel $\lambda \neq 0$ tel que : $u_n \underset{+\infty}{\sim} \lambda v_n$.
  Alors les séries $\sum u_n$ et $\sum v_n$ sont de même nature.

  ✓ Il est important de se rappeler que cette règle ne s'applique qu'à des séries de signe constant (au moins à partir d'un certain rang).
- **Règle de comparaison n° 4**
  Soit (u_n) une suite à valeurs réelles et (v_n) une suite de nombres réels *positifs*.
  On suppose que : $u_n \underset{+\infty}{=} O(v_n)$ (ou que $u_n \underset{+\infty}{=} o(v_n)$).
  Alors, si la série $\sum v_n$ converge, la série $\sum u_n$ converge (absolument).

  ✓ On fera très attention là encore lors de l'utilisation de cette règle : la série $\sum u_n$ est à termes réels quelconques, mais il est indispensable que la série $\sum v_n$ à laquelle on la compare soit, elle, à *termes réels positifs*.

- **Règle de d'Alembert**

  Soit $\sum u_n$ une série à termes réels strictement positifs On suppose que $\lim\limits_{n\to+\infty}\left(\dfrac{u_{n+1}}{u_n}\right) = \ell$ existe dans $\overline{\mathbb{R}}$ ($\ell \in [0\,;+\infty]$).

  – Si $\ell < 1$, la série $\sum u_n$ converge.

  – Si $\ell > 1$, la série $\sum u_n$ diverge.

  ✓ Lorsque $\ell = 1$, on ne peut rien dire a priori . Exemples : $\sum \frac{1}{n}$ et $\sum \frac{1}{n^2}$.

[S11.7] Comparaison à une intégrale

- Soit f une fonction continue par morceaux sur un intervalle de la forme $[n_0\,;+\infty[$ ($n_0 \in \mathbb{N}$), *à valeurs réelles positives* et *décroissante*. Alors :

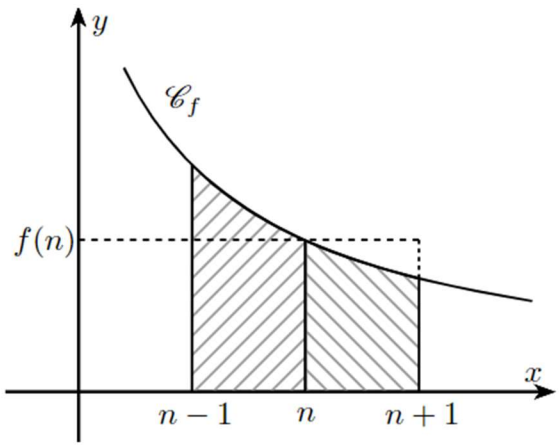
$$\text{la série } \sum_{n\geqslant n_0} f(n) \text{ converge} \iff \int_{n_0}^{+\infty} f \text{ existe}\,.$$

  ✓ Il est important de retenir non seulement ce théorème, mais aussi et surtout sa démonstration. En effet, la méthode consistant à encadrer le terme général $f(n)$ par des intégrales de f sur des segments permet d'obtenir plus généralement des encadrements des sommes partielles ou du reste de la série lorsqu'elle converge.

Il faut donc retenir (ou savoir retrouver) l'encadrement suivant, valable pour tout $n \geqslant n_0 + 1$:

$$\int_n^{n+1} f(t)\,\mathrm{d}t \leqslant f(n) \leqslant \int_{n-1}^{n} f(t)\,\mathrm{d}t\,.$$

Il est illustré par la figure ci-contre.



- **Application : les séries de Riemann**

  La série $\sum\limits_{n\in\mathbb{N}^*} \dfrac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

- **Généralisation**

  Soit f une fonction continue par morceaux sur un intervalle de la forme $[n_0\,;+\infty[$ ($n_0 \in \mathbb{N}$), *à valeurs réelles positives* et *décroissante*. Alors la série de terme général :

$$w_n = \int_{n-1}^{n} f(t)\,\mathrm{d}t - f(n) \quad (n \geqslant n_0 + 1)$$

  est convergente.

- **Application : la constante d'Euler**
  Il existe un réel γ tel que, pour $n \in \mathbb{N}^*$:

$$1 + \frac{1}{2} + \cdots + \frac{1}{n} \underset{n\to+\infty}{=} \ln n + \gamma + o(1)\,.$$

[S11.8] Séries géométriques

Soit $(\mathcal{A}, +, .,\times)$ une algèbre normée de dimension finie.

Pour tout $x \in \mathcal{A}$ tel que $\|x\| < 1$, la série géométrique $\sum_{n\in\mathbb{N}} x^n$ est absolument convergente (donc convergente). De plus, $1_\mathcal{A} - x$ est inversible, et :

$$(1_\mathcal{A} - x)^{-1} = \sum_{n=0}^{+\infty} x^n\,.$$

Enfin, l'application $x \mapsto (1_\mathcal{A} - x)^{-1}$ est continue sur la boule unité ouverte $\mathcal{B}(0_\mathcal{A},1)$.

[S11.9] Série exponentielle

Soit $(\mathcal{A}, +, .,\times)$ une algèbre normée de dimension finie.

Pour tout $x \in \mathcal{A}$, la série $\sum_{n\in\mathbb{N}} \frac{1}{n!}\, x^n$ est absolument convergente (donc convergente).

On pose alors :

$$\forall\, x \in \mathcal{A},\ \exp(x) = \sum_{n=0}^{+\infty} \frac{1}{n!}\, x^n.$$

La fonction exponentielle $x \mapsto \exp(x)$ est continue sur $\mathcal{A}$.

> ✓ Lorsque x est un nombre réel, l'inégalité de Taylor-Lagrange permet de démontrer que cette définition coïncide avec celle de la fonction exponentielle que vous connaissiez déjà.

Exemple : on peut munir l'algèbre $\mathcal{M}_p(\mathbb{K})$ ($p \in \mathbb{N}^*$) d'une *norme matricielle*, c'est-à-dire qui vérifie :

$$\forall\, A,\, B \in \mathcal{M}_p(\mathbb{K}),\ \|AB\| \leqslant \|A\|\,\|B\|\,.$$

Le résultat précédent permet alors de définir l'exponentielle d'une matrice carrée, et plus généralement l'exponentielle d'un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie.

[S11.10] Produit de Cauchy

Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites à valeurs dans une algèbre normée $(\mathcal{A}, +, .,\times)$ de dimension finie.

On appelle série produit de Cauchy des séries de terme général u_n d'une part et de terme général v_n d'autre part, la série de terme général w_n avec :

$$\forall\, n \in \mathbb{N},\ w_n = \sum_{k=0}^{n} u_k \times v_{n-k} = \sum_{\substack{p+q=n \\ p,q\in\mathbb{N}}} u_p \times v_q\,.$$

Si $\sum u_n$ et $\sum v_n$ sont absolument convergentes, alors $\sum w_n$ est absolument convergente, et de plus :

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n\right)\left(\sum_{n=0}^{+\infty} v_n\right).$$

Ce théorème permet de démontrer que si x et y sont deux éléments *qui commutent* d'une algèbre normée $(\mathcal{A}, +, ., \times)$ de dimension finie, on a :

$$\exp(x+y) = \exp(x) \times \exp(y).$$

On en déduit que pour tout $x \in \mathcal{A}$, $\exp(x)$ est inversible, d'inverse $\exp(-x)$.

[S11.11] Séries alternées de nombres réels

Une série à termes réels $\sum u_n$ est dite alternée si la suite $\left((-1)^n u_n\right)$ est de signe constant.

- **Critère spécial des séries alternées, ou critère de Leibniz**
  Soit $\sum u_n$ une série alternée. On suppose que :
  - la suite $(|u_n|)$ est décroissante ;
  - $\lim\limits_{n\to+\infty} u_n = 0$.

  Alors la série $\sum u_n$ converge.

  ✓ Il est important de noter que ce critère ne donne qu'une condition *suffisante* de convergence pour une série alternée. Par exemple, la série de terme général $u_n = \dfrac{(-1)^n}{n+(-1)^n}$ $(n \geqslant 2)$ est convergente, (on le montre en effectuant un développement limité), cependant elle ne vérifie pas les hypothèses du critère de Leibniz (la suite $(|u_n|)$ n'étant pas décroissante).

- **Résultats complémentaires importants**
  Soit $\sum\limits_{n\in\mathbb{N}} u_n$ une série alternée *qui vérifie les hypothèses du critère spécial*, et soit $S = \sum\limits_{n=0}^{+\infty} u_n$ sa somme. On a les résultats suivants.

  (i) S est comprise entre deux sommes partielles d'indices consécutifs.

  (ii) S est du signe de u_0, et $|S| \leqslant |u_0|$.

  (iii) Si l'on note $R_n = \sum\limits_{k=n+1}^{+\infty} u_k$ le reste d'ordre n, alors R_n est du signe de u_{n+1} et $|R_n| \leqslant |u_{n+1}|$.

[S11.12] Sommation des relations de comparaison

- **Cas des séries convergentes**

  Soit $\sum_{n\in\mathbb{N}} u_n$ une série à termes réels et $\sum_{n\in\mathbb{N}} v_n$ une série *à termes réels positifs* (au moins à partir d'un certain rang) et *convergente*.

  On suppose que, lorsque $n \to +\infty$, $u_n = O(v_n)$ (respectivement $u_n = o(v_n)$, respectivement $u_n \sim v_n$).

  Alors :

  – la série $\sum_{n\in\mathbb{N}} u_n$ est absolument convergente ;

  – si l'on note, pour tout entier n, $r_n = \sum_{k=n+1}^{+\infty} u_k$ et $r'_n = \sum_{k=n+1}^{+\infty} v_k$, on a :

  $$r_n = O(r'_n) \text{ (respectivement } r_n = o(r'_n), \text{ respectivement } r_n \sim r'_n).$$

- **Cas des séries divergentes**

  Soit $\sum_{n\in\mathbb{N}} u_n$ une série à termes réels et $\sum_{n\in\mathbb{N}} v_n$ une série *à termes réels positifs* (au moins à partir d'un certain rang), toutes deux *divergentes*.

  On suppose que, lorsque $n \to +\infty$, $u_n = O(v_n)$ (respectivement $u_n = o(v_n)$, respectivement $u_n \sim v_n$).

  Alors, si l'on note, pour tout entier n, $S_n = \sum_{k=0}^{n} u_k$ et $S'_n = \sum_{k=0}^{n} v_k$, on a :

  $$S_n = O(S'_n) \text{ (respectivement } S_n = o(S'_n), \text{ respectivement } S_n \sim S'_n).$$

[S11.13] Ensembles dénombrables

- Deux ensembles E et F sont dits équipotents s'il existe une bijection de E sur F. C'est une relation d'équivalence.

  Dans le cas où E et F sont finis, ils sont équipotents si et seulement si ils ont même cardinal.

- Un ensemble est dit dénombrable s'il est équipotent à $\mathbb{N}$.

  E est dénombrable si et seulement si on peut le décrire sous la forme

  $$E = \{x_n,\, n \in \mathbb{N}\}\ .$$

  Toute partie infinie d'un ensemble dénombrable est dénombrable.

- Un ensemble E est dit au plus dénombrable s'il est équipotent à une partie de $\mathbb{N}$. Cela équivaut à dire qu'il est soit fini, soit dénombrable.

  E est au plus dénombrable si et seulement si on peut le décrire sous la forme

  $$E = \{x_i,\, i \in I\} \quad \text{où} \quad I \text{ est une partie de } \mathbb{N}\,.$$

- **Ensembles usuels**

  . $\mathbb{Z}$ est dénombrable.

- $\mathbb{N} \times \mathbb{N}$ est dénombrable.
- Plus généralement, si E et F sont dénombrables, leur produit $E \times F$ l'est aussi.
- $\mathbb{Q}$ est dénombrable.
- Une réunion finie ou dénombrable d'ensembles dénombrables est un ensemble dénombrable.
- L'ensemble $\mathcal{P}(\mathbb{N})$ des parties de $\mathbb{N}$ n'est pas dénombrable.
- L'ensemble $\{0,1\}^{\mathbb{N}}$ des suites à valeurs dans $\{0,1\}$ n'est pas dénombrable.
- $\mathbb{R}$ n'est pas dénombrable.

[S11.14] Familles sommables

- **Famille sommable de réels positifs**

  Soit I un ensemble dénombrable d'indices, et soit $(u_i)_{i\in I}$ une famille de réels *positifs ou nuls.*

  On dit que la famille est sommable s'il existe un réel positif M tel que

  $$\sum_{i\in J} u_i \leqslant M$$

  pour toute partie finie J incluse dans I.

  Dans ces conditions, la borne supérieure de l'ensemble des sommes finies $\sum_{i\in J} u_i$ (J parcourant l'ensemble de toutes les parties finies de I) est appelée la somme de la famille $(u_i)_{i\in I}$, et est notée $\sum_{i\in I} u_i$.

  Dans le cas où la famille n'est pas sommable, on écrira $\sum_{i\in I} u_i = +\infty$.

- **Famille sommable de nombres réels**

  Soit $(u_i)_{i\in I}$ une famille de nombres réels quelconque, indexée par un ensemble dénombrable I.

  La famille $(u_i)_{i\in I}$ est dite sommable lorsque la famille $(|u_i|)_{i\in I}$ est une famille sommable de réels positifs.

  Rappelons que, pour tout nombre réel x, on pose :

  $$x^+ = \max(x,0) \quad \text{et} \quad x^- = \max(-x,0)\,,$$

  de sorte que x^+ et x^- sont des réels positifs tels que :

  $$x = x^+ - x^- \quad \text{et} \quad |x| = x^+ + x^-\,.$$

  Pour que la famille de nombres réels $(u_i)_{i\in I}$ soit sommable, il faut et il suffit que les deux familles de nombres réels positifs $(u_i^+)_{i\in I}$ et $(u_i^-)_{i\in I}$ le soient.

  On posera alors, dans ce cas : $\sum_{i\in I} u_i = \sum_{i\in I} u_i^+ - \sum_{i\in I} u_i^-$.

- **Famille sommable de nombres complexes**

  Soit $(u_i)_{i\in I}$ une famille de nombres complexes, indexée par un ensemble dénombrable I.

  La famille $(u_i)_{i\in I}$ est dite sommable lorsque la famille $(|u_i|)_{i\in I}$ est une famille sommable de réels positifs.

  Pour que la famille de nombres complexes $(u_i)_{i\in I}$ soit sommable, il faut et il suffit que les deux familles de nombres réels $\big(\mathcal{R}e\,(u_i)\big)_{i\in I}$ et $\big(\mathcal{I}m\,(u_i)\big)_{i\in I}$ le soient.

  On posera alors, dans ce cas : $\displaystyle\sum_{i\in I} u_i = \sum_{i\in I} \mathcal{R}e\,(u_i) + \mathrm{i}\sum_{i\in I} \mathcal{I}m\,(u_i)$.

- **Linéarité de la somme**

  Soient $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ deux familles sommables de nombres complexes, et λ, μ deux nombres complexes.

  Alors la famille $(\lambda u_i + \mu v_i)_{i\in I}$ est sommable et :

$$\sum_{i\in I} \lambda u_i + \mu v_i = \lambda \sum_{i\in I} u_i + \mu \sum_{i\in I} v_i\,.$$

- **Lien avec les séries**

  Dans le cas où $I = \mathbb{N}$ on a le résultat suivant.

  Une suite $(u_n)_{n\in\mathbb{N}}$ est sommable si et seulement si la série $\sum u_n$ est absolument convergente, et, dans ce cas, on a :

$$\sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n\,.$$

- **Commutativité généralisée**

  Pour qu'une famille dénombrable $(u_i)_{i\in I}$ de nombres complexes soit sommable, il faut et il suffit que pour une/toute bijection φ de $\mathbb{N}$ sur I la série de terme général $u_{\varphi(n)}$ soit absolument convergente. Et dans ce cas l'on a :

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty} u_{\varphi(n)}\,,$$

  la somme ne dépendant pas de la bijection $\varphi\colon \mathbb{N} \to I$ choisie.

- **Associativité généralisée, ou sommation par paquets**

  Soit $(u_i)_{i\in I}$ une famille dénombrable de nombres complexes, et $(I_n)_{n\in N}$ une partition dénombrable de I.

  La famille $(u_i)_{i\in I}$ est sommable si et seulement si :

  (i) chaque sous-famille $(u_i)_{i\in I_n}$ est sommable ;

  (ii) la série $\displaystyle\sum_{n\in\mathbb{N}} \left(\sum_{i\in I_n} |u_i|\right)$ est convergente.

  Dans ce cas, l'on a :

$$\sum_{n=0}^{+\infty} \left(\sum_{i\in I_n} u_i\right) = \sum_{i\in I} u_i\,.$$

- **Suites doubles sommables**

  On prend ici $I = \mathbb{N} \times \mathbb{N}$. Une famille $(u_{p,q})_{(p,q)\in\mathbb{N}^2}$ est appelée une suite double.

  Théorème de Fubini

  La suite double $(u_{p,q})_{(p,q)\in\mathbb{N}^2}$ est sommable si et seulement si les deux conditions suivantes sont réalisées :

  – pour tout q fixé dans $\mathbb{N}$, la série $\sum_{p\in\mathbb{N}} u_{p,q}$ est absolument convergente ;

  – la série de terme général $S_q = \sum_{p=0}^{+\infty} |u_{p,q}|$ est convergente.

  Dans ces conditions :

  – pour tout p fixé dans $\mathbb{N}$, la série $\sum_{q\in\mathbb{N}} u_{p,q}$ est absolument convergente ;

  – la série de terme général $S'_p = \sum_{q=0}^{+\infty} |u_{p,q}|$ est convergente ;

  – on a les égalités :

$$\sum_{p=0}^{+\infty} S'_p = \sum_{p=0}^{+\infty} \left(\sum_{q=0}^{+\infty} |u_{p,q}| \right) = \sum_{q=0}^{+\infty} S_q = \sum_{q=0}^{+\infty} \left(\sum_{p=0}^{+\infty} |u_{p,q}| \right)$$

$$\sum_{(p,q)\in\mathbb{N}^2} u_{pq} = \sum_{p=0}^{+\infty} \left(\sum_{q=0}^{+\infty} u_{p,q} \right) = \sum_{q=0}^{+\infty} \left(\sum_{p=0}^{+\infty} u_{p,q} \right).$$

- **Suite double produit**

  Un cas particulier important et fréquent d'application du théorème de Fubini est le suivant.

  Soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites de nombres complexes.
  Pour que la suite double $(a_p b_q)_{(p,q)\in\mathbb{N}^2}$ soit sommable il faut et il suffit que les séries $\sum_{p\in\mathbb{N}} a_p$ et $\sum_{q\in\mathbb{N}} b_q$ soient absolument convergentes, et on a alors :

$$\sum_{p,q\in\mathbb{N}} a_p b_q = \left(\sum_{p=0}^{+\infty} a_p \right) \left(\sum_{q=0}^{+\infty} b_q \right).$$

Thème 12 - Suites et séries de fonctions

Dans ce chapitre, les fonctions sont définies sur une partie non vide A d'un $\mathbb{K}$-espace vectoriel normé de dimension finie E et à valeurs dans un $\mathbb{K}$-espace vectoriel normé de dimension finie F.

Les normes sur ces espaces étant équivalentes, les notions introduites ne dépendent pas des normes choisies.

[S12.1] Convergence d'une suite de fonctions

- Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications définies sur A, à valeurs dans F.
  On dit que cette suite converge simplement sur A si et seulement si :
  $$\text{pour tout } x \in A,\ \lim_{n\to+\infty} f_n(x) \text{ existe (dans } F).$$
  Dans ce cas, on peut définir une application $f\colon A \longrightarrow F$ par :
  $$\forall\, x \in A,\ f(x) = \lim_{n\to+\infty} f_n(x)\,.$$
  f s'appelle la limite simple de la suite $(f_n)_{n\in\mathbb{N}}$.
- On dit que la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur A vers f si, à partir d'un certain rang, la suite $(f_n - f)$ est bornée sur A et si :
  $$\lim_{n\to+\infty} \|f_n - f\|_\infty^A = 0\,,$$
  où $\|f_n - f\|_\infty^A$ désigne $\sup\limits_{x\in A} \|f_n(x) - f(x)\|_F$ (norme de la convergence uniforme).

  ✓ Lorsque l'on étudie la convergence uniforme, il est indispensable de préciser sur quelle partie elle a lieu ; la phrase « $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f » ne veut rien dire si l'on ne précise pas « sur A ».

[S12.2] Théorème de la double limite

Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications de A dans F, qui converge *uniformément* sur A vers une application $f\colon A \to F$.

Soit a un point adhérent à A. On suppose que, pour tout entier n (au moins à partir d'un certain rang), la limite $\lim\limits_{\substack{x\to a\\ x\in A}} f_n(x) = \ell_n$ existe.

Alors la suite $(\ell_n)_{n\in\mathbb{N}}$ converge vers un élément $\ell \in F$, et de plus : $\lim\limits_{\substack{x\to a\\ x\in A}} f(x) = \ell$.

Autrement dit, en abrégé : $\lim\limits_{x\to a}\left(\lim\limits_{n\to+\infty} f_n(x)\right) = \lim\limits_{n\to+\infty}\left(\lim\limits_{x\to a} f_n(x)\right)$.

✓ Lorsque f est définie sur un intervalle de $\mathbb{R}$ qui est un voisinage de $\pm\infty$, le résultat de ce théorème peut s'étendre aux cas $a = \pm\infty$.

[S12.3] Continuité de la fonction limite

- Soit $a \in A$ et $(f_n)_{n\in\mathbb{N}}$ une suite d'applications de A dans F, continues en a, qui converge vers une application $f\colon A \to F$, *la convergence étant uniforme sur un voisinage de a.*

  Alors f est continue en a

- Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications continues de A dans F, qui converge vers une application $f\colon A \to F$, *la convergence étant uniforme au voisinage de tout point de A.* Alors f est continue sur A

  ✓ Ce théorème peut dans certains cas servir à démontrer qu'il n'y a pas convergence uniforme, en utilisant un raisonnement par l'absurde.

[S12.4] Intégration sur un segment

- Soit (f_n) une suite d'applications continues définies sur un intervalle I de $\mathbb{R}$ et à valeurs dans F, et soit $a \in I$. On suppose que la suite (f_n) converge simplement sur I vers une application $f\colon I \to F$, *la convergence étant uniforme sur tout segment inclus dans I.*

  Pour tout $n \in \mathbb{N}$ et tout $x \in I$ on pose :

$$g_n(x) = \int_a^x f_n(t)\,\mathrm{d}t \quad \text{et} \quad g(x) = \int_a^x f(t)\,\mathrm{d}t\,.$$

  Alors la suite (g_n) converge uniformément vers g sur tout segment de I.

- **Cas particulier**

  Soit (f_n) une suite d'applications continues définies sur un segment $[a\,;b]$ à valeurs dans F, qui converge uniformément sur $[a\,;b]$ vers $f\colon [a\,;b] \to F$. Alors :

$$\lim_{n\to+\infty} \int_a^b f_n(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t\,.$$

[S12.5] Dérivation

- **Dérivabilité de la limite d'une suite de fonctions de classe $\mathscr{C}^1$**

  Soit $(f_n)_{n\in\mathbb{N}}$ une suite d'applications définies sur un intervalle I de $\mathbb{R}$, à valeurs dans F. On suppose que :

  – les f_n sont de classe $\mathscr{C}^1$ sur I ;

  – la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur I vers une fonction f ;

  – la suite des dérivées $(f'_n)_{n\in\mathbb{N}}$ converge vers une application $g\colon I \to F$, *la convergence étant uniforme sur tout segment inclus dans I.*

  Alors, la fonction f est de classe $\mathscr{C}^1$ sur I, la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout segment inclus dans I, et, pour tout $x \in I$, $f'(x) = g(x)$ (soit, en abrégé, $(\lim f_n)' = \lim f'_n$).

- **Dérivabilité de la limite d'une suite de fonctions de classe $\mathscr{C}^k$**

  Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de classe $\mathscr{C}^k$ ($k \in \mathbb{N}^*$) sur un intervalle I de $\mathbb{R}$, à valeurs dans F. On suppose que :

  – pour tout $j \in [\![0\,;k-1]\!]$, la suite de fonctions $(f_n^{(j)})_{n\in\mathbb{N}}$ converge simplement sur I ;

  – la suite de fonctions $(f_n^{(k)})_{n\in\mathbb{N}}$ converge simplement sur I vers une fonction g, *la convergence étant uniforme sur tout segment inclus dans* I.

  Alors, la fonction $f = \lim\limits_{n\to+\infty} f_n$ est de classe $\mathscr{C}^k$ sur I, on a $f^{(k)} = g$, et pour $j \in [\![0\,;k]\!]$, chaque suite $(f_n^{(j)})_{n\in\mathbb{N}}$ converge uniformément vers $f^{(j)}$ sur tout segment inclus dans I.

  ✓ Pour les deux théorèmes précédents, plutôt que de démontrer la convergence uniforme sur *tout segment* de I, on peut se contenter de montrer la convergence uniforme sur tous les intervalles d'une famille (J_α) telle que $\bigcup\limits_\alpha J_\alpha = I$.

[S12.6] Approximation uniforme

- **Approximation uniforme d'une fonction continue par morceaux par des fonctions en escalier**

  Toute fonction continue par morceaux sur un segment $[a\,;b]$ à valeurs dans un espace vectoriel normé F est limite uniforme sur $[a\,;b]$ d'une suite de fonctions en escalier sur $[a\,;b]$.

- **Théorème de Weierstrass**

  Toute fonction continue sur un segment $[a\,;b]$ à valeurs dans $\mathbb{C}$ est limite uniforme sur $[a\,;b]$ d'une suite de fonctions polynomiales sur $[a\,;b]$.

[S12.7] Convergence d'une série de fonctions

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'applications définies sur une partie A d'un $\mathbb{K}$-espace vectoriel normé E, à valeurs dans un $\mathbb{K}$-espace vectoriel normé F.
On peut alors considérer la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ définies par :

$$\forall\, x \in A,\ S_n(x) = \sum_{k=0}^{n} u_k(x)\,.$$

Étudier la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$, c'est étudier la suite de fonctions $(S_n)_{n\in\mathbb{N}}$.

- **Convergence simple**

  On dit que la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ converge simplement sur A s'il existe une application $S\colon A \to F$ telle que la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ converge simplement sur A vers S.

Cela signifie donc que, pour tout $x \in A$, la série $\sum_{n\in\mathbb{N}} u_n(x)$, à valeurs dans F, converge et que $S(x) = \sum_{n=0}^{+\infty} u_n(x)$.

S s'appelle alors la somme de la série de fonctions $\sum_{n\in\mathbb{N}} u_n$. On définit également le reste d'ordre n : $R_n = S - S_n = \sum_{k=n+1}^{+\infty} u_k$. La suite de fonctions $(R_n)_{n\in\mathbb{N}}$ converge simplement sur A vers la fonction nulle.

- **Convergence uniforme**

  On dit que la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge uniformément sur A s'il existe une application $S\colon A \to F$ telle que la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ converge uniformément sur A vers S.

  Cela équivaut à dire que la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge simplement sur A et que la suite des restes $(R_n)_{n\in\mathbb{N}}$ converge *uniformément* sur A vers la fonction nulle (c'est-à-dire $\lim_{n\to+\infty} \|R_n\|_\infty^A = 0$).

  ✓ La convergence uniforme sur I de la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ implique nécessairement la convergence uniforme de la suite (u_n) vers la fonction nulle sur I (puisque $u_n = R_{n-1} - R_n$).
  En raisonnant par l'absurde, cette remarque peut être utilisée pour montrer qu'une série de fonctions ne converge pas uniformément sur I : il suffit pour cela de trouver une suite (x_n) d'éléments de I telle que la suite $\big(u_n(x_n)\big)$ ne converge pas vers 0.

- **Convergence normale**

  On dit que la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge normalement sur A si :

  – les fonctions u_n sont bornées sur A (au moins à partir d'un certain rang),

  – et la série numérique $\sum_{n\geqslant 0} \|u_n\|_\infty^A$ est convergente (en notant toujours $\|u_n\|_\infty^A = \sup_{x\in A} \|u_n(x)\|_F$).

  Si la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ est *normalement convergente* sur A, alors :

  (i) pour tout $x \in A$, la série $\sum_{n\in\mathbb{N}} u_n(x)$ est absolument convergente dans F (donc convergente puisque F est de dimension finie) ;

  (ii) la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ est uniformément convergente sur A.

[S12.8] Convergence uniforme d'une série de fonctions et limite

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur A à valeurs dans F, et soit $a \in \overline{A}$.

On suppose que, pour tout entier n, la limite $\lim\limits_{\substack{x\to a\\ x\in A}} u_n(x) = \ell_n$ existe dans F, et que la série $\sum_{n\in\mathbb{N}} u_n$ est *uniformément convergente sur* A. Notons $S = \sum_{n=0}^{+\infty} u_n$. Alors :

$$\text{la série } \sum_{n\in\mathbb{N}} \ell_n \text{ converge et } \lim_{\substack{x\to a\\ x\in A}} S(x) = \sum_{n=0}^{+\infty} \ell_n$$

(soit, en abrégé : $\lim\limits_a \left(\sum_{n=0}^{+\infty} u_n\right) = \sum_{n=0}^{+\infty} \lim\limits_a u_n$).

Ce résultat porte le nom de *théorème de la double limite.*

[S12.9] Convergence uniforme d'une série de fonctions et continuité

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur A, à valeurs dans F.
Si les u_n sont continues sur A et si la série *converge uniformément au voisinage de tout point de* A, alors sa somme S est continue sur A.

[S12.10] Convergence d'une série de fonctions et intégration sur un segment

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur *un segment* $[a\,;b] \subset \mathbb{R}$, à valeurs dans F. On suppose que les u_n sont continues sur $[a\,;b]$, et que la série $\sum_{n\in\mathbb{N}} u_n$ converge *uniformément sur* $[a\,;b]$. Notons $S = \sum_{n=0}^{+\infty} u_n$.

Alors S est continue sur $[a\,;b]$, la série $\sum_{n\in\mathbb{N}} \int_a^b u_n(t)\,\mathrm{d}t$ converge, et :

$$\int_a^b S(t)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_a^b u_n(t)\,\mathrm{d}t\,.$$

Remarque : le théorème s'applique également lorsque les u_n sont seulement *continues par morceaux*, mais il faut alors vérifier la continuité par morceaux de S, celle-ci n'étant plus assurée par la convergence uniforme.

[S12.11] Convergence d'une série de fonctions et dérivation

- Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur un intervalle I de $\mathbb{R}$, à valeurs dans F. On suppose que :

– les u_n sont de classe $\mathscr{C}^1$ sur I ;

– la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge simplement sur I ; on notera S sa somme ;

– la série de fonctions $\sum_{n\in\mathbb{N}} u'_n$ converge simplement sur I, *la convergence étant uniforme sur tout segment inclus dans* I.

Alors :

– la fonction S est de classe $\mathscr{C}^1$ sur I ;

– la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge uniformément sur tout segment de I ;

– pour tout $x \in I$, on a : $S'(x) = \sum_{n=0}^{+\infty} u'_n(x)$.

- Comme pour les suites de fonctions, on a également un énoncé pour démontrer que la somme d'une série de fonctions est de classe $\mathscr{C}^k$.

Soit $\sum_{n\in\mathbb{N}} u_n$ une série de fonctions définies sur un intervalle I, à valeurs dans F.
On suppose que :

– les u_n sont de classe $\mathscr{C}^k$ sur I ;

– $\forall\, j \in [\![0\,;k-1]\!]$, la série de fonctions $\sum_{n\in\mathbb{N}} u_n^{(j)}$ converge simplement sur I ;

– la série de fonctions $\sum_{n\in\mathbb{N}} u_n^{(k)}$ converge simplement sur I, *la convergence étant uniforme sur tout segment inclus dans* I.

Alors la fonction somme S est de classe $\mathscr{C}^k$ sur I, chaque série $\sum u_n^{(j)}$ avec $j \in [\![0\,;k]\!]$ converge uniformément vers $S^{(j)}$ sur tout segment de I et :

$$\forall\, j \in [\![0\,;k]\!],\ \forall\, x \in I,\ S^{(j)}(x) = \sum_{n=0}^{+\infty} u_n^{(j)}(x)\,.$$

✓ Pour les deux théorèmes précédents, plutôt que de démontrer la convergence uniforme sur *tout segment* de I, on peut se contenter de montrer la convergence uniforme sur tous les intervalles d'une famille (J_α) telle que $\bigcup_\alpha J_\alpha = I$.

Thème 13 - Séries entières

[S13.1] Rayon de convergence d'une série entière

- Une série entière est une série de fonctions de la forme :

$$\sum_{n\in\mathbb{N}} a_n z^n \,,$$

où z est la variable complexe et $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes.

- **Lemme d'Abel**

S'il existe un complexe non nul z_0 tel que la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ soit bornée, alors pour tout nombre complexe z tel que $|z| < |z_0|$, la série numérique $\sum_{n\in\mathbb{N}} a_n z^n$ est absolument convergente.

- **Définition du rayon de convergence**

Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière. L'ensemble des réels positifs r tels que la suite $(a_n r^n)_{n\in\mathbb{N}}$ soit bornée est un intervalle contenant 0.
On appelle alors rayon de convergence R de cette série entière la borne supérieure (dans $\overline{\mathbb{R}}$) de cet intervalle :

$$R = \sup\{r \in \mathbb{R}_+ \mid (a_n r^n) \text{ bornée}\} \quad (R \in \mathbb{R}_+ \cup \{+\infty\})\,.$$

- **Disque de convergence**

Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière de rayon de convergence R.

- Si $R > 0$, alors pour tout z tel que $|z| < R$, la série $\sum_{n\in\mathbb{N}} a_n z^n$ est absolument convergente.
- Si $R < +\infty$ alors pour tout z tel que $|z| > R$, la série $\sum_{n\in\mathbb{N}} a_n z^n$ est (grossièrement) divergente.
- Si $R = 0$, la série ne converge que pour $z = 0$.
- Si $R = +\infty$, la série est convergente (absolument) pour tout $z \in \mathbb{C}$.

La boule ouverte $\mathcal{B}(0,R)$ dans $\mathbb{C}$ s'appelle le disque de convergence.

> ✓ On ne peut rien dire dans le cas général du comportement de la série entière sur le cercle d'incertitude $\{z \in \mathbb{C} \mid |z| = R\}$.

- **Autres caractérisations du rayon de convergence**

Le rayon de convergence R d'une série entière $\sum_{n\in\mathbb{N}} a_n z^n$ peut être défini par l'une des propriétés suivantes, toutes équivalentes :

- $R = \sup\left\{ |z|\,,\ z \in \mathbb{C} \,\middle|\, \text{la série } \sum_{n\in\mathbb{N}} a_n z^n \text{ est absolument convergente} \right\}$;

– $R = \sup\left\{ |z|\,,\ z \in \mathbb{C} \,\middle|\, \text{la série } \sum_{n\in\mathbb{N}} a_n z^n \text{ est convergente} \right\}$;

– $R = \sup\left\{ |z|\,,\ z \in \mathbb{C} \,\middle|\, \lim_{n\to+\infty} a_n z^n = 0 \right\}$;

– $R = \sup\left\{ |z|\,,\ z \in \mathbb{C} \,\middle|\, \text{la suite } (a_n z^n) \text{ est bornée} \right\}$ (définition).

- **Utilisation des règles de comparaison**

  Soient $\sum_{n\in\mathbb{N}} a_n z^n$ et $\sum_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergence respectifs R_a et R_b.

  – Si $a_n \underset{n\to+\infty}{\sim} b_n$, alors $R_a = R_b$.

  – Si $|a_n| \leqslant |b_n|$ (au moins à partir d'un certain rang), alors $R_a \geqslant R_b$.

  – Si $a_n = O(b_n)$ (ou $a_n = o(b_n)$), alors $R_a \geqslant R_b$.

- **Critère de d'Alembert**

  Si la suite $(a_n)_{n\in\mathbb{N}}$ *ne s'annule pas* (au moins à partir d'un certain rang), et si la suite $\left(\left|\frac{a_{n+1}}{a_n}\right|\right)_{n\in\mathbb{N}}$ possède une limite $\ell \in \overline{\mathbb{R}}$ lorsque $n \to +\infty$, alors la série entière $\sum_{n\in\mathbb{N}} a_n z^n$ a pour rayon de convergence : $R = \frac{1}{\ell}$.

  ✓ On fera particulièrement attention à ne pas utiliser ce critère dans le cas où la série entière peut avoir des coefficients nuls. Dans ce cas, il convient d'utiliser la règle de d'Alembert pour les séries numériques (cf. **[S11.6]** page 122) pour étudier la convergence de la série $\sum |a_n z^n|$.

[S13.2] Opérations sur les séries entières

- **Somme de deux séries entières**

  Soient $\sum_{n\in\mathbb{N}} a_n z^n$ et $\sum_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergence respectifs R_a et R_b.

  – Si $R_a \neq R_b$, la série entière $\sum_{n\in\mathbb{N}} (a_n + b_n) z^n$ a pour rayon de convergence $R = \min(R_a, R_b)$.

  – Si $R_a = R_b$, la série entière $\sum_{n\in\mathbb{N}} (a_n + b_n) z^n$ a un rayon de convergence $R \geqslant R_a$.

  – Dans les deux cas on a, pour tout z tel que $|z| < \min(R_a, R_b)$:

$$\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n .$$

- **Produit de Cauchy de deux séries entières**

  Soient $\sum_{n\in\mathbb{N}} a_n z^n$ et $\sum_{n\in\mathbb{N}} b_n z^n$ deux séries entières de rayons de convergence respectifs R_a et R_b.

  Posons, pour tout $n \in \mathbb{N}$: $c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{\substack{p,q\in\mathbb{N} \\ p+q=n}} a_p b_q$.

– Le rayon de convergence R_c de la série entière $\sum_{n\in\mathbb{N}} c_n z^n$ est tel que $R_c \geqslant \min(R_a, R_b)$.
– De plus, pour tout complexe z tel que $|z| < \min(R_a, R_b)$, on a :

$$\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right)\left(\sum_{n=0}^{+\infty} b_n z^n\right).$$

[S13.3] Convergence normale d'une série entière

La série entière $\sum_{n\in\mathbb{N}} a_n z^n$ est normalement convergente sur tout *compact* inclus dans son disque ouvert de convergence.

> ✓ Il n'y a pas nécessairement convergence normale ni même convergence uniforme de la série entière sur *tout* le disque de convergence. Le contre-exemple qui suit est classique, et doit être su.
>
> *La série entière de la variable réelle* $\sum_{n\in\mathbb{N}^*} (-1)^{n-1}\frac{x^n}{n}$ *converge simplement vers* $x \mapsto \ln(1+x)$ *sur* $]-1\,;1]$.
> *Il n'y a pas convergence normale sur* $]-1\,;1[$ *(car* $\|u_n\|_\infty^{]-1;1[} = \frac{1}{n}$*).*
> *Il n'y a pas convergence uniforme sur* $]-1\,;1[$ *(car le reste* $\sum_{k=n+1}^{+\infty} (-1)^{k-1}\frac{x^k}{k}$ *n'est pas borné au voisinage de* -1*).*
> *Il y a cependant convergence uniforme sur tout segment* $[a\,;1]$ *avec* $-1 < a < 1$ *(cela se montre en utilisant la majoration du reste d'une série alternée).*

[S13.4] Continuité de la somme d'une série entière

La fonction somme d'une série entière $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ est continue sur son disque ouvert de convergence.

Conséquence
Si f est la somme d'une série entière de la variable réelle x, f admet au voisinage de 0 un développement limité à tout ordre $p \in \mathbb{N}^*$, obtenu par troncature de la série entière : $f(x) = \sum_{n=0}^{p} a_n x^n + O\left(x^{p+1}\right)$.

[S13.5] Dérivation

- **Séries dérivée et primitive**

  Soit $\sum_{n\in\mathbb{N}} a_n z^n$ une série entière.

  – On appelle série dérivée de cette série la série entière :

$$\sum_{n\geqslant 1} n a_n z^{n-1} = \sum_{n\geqslant 0} (n+1) a_{n+1} z^n.$$

– On appelle série primitive de cette série la série entière :

$$\sum_{n\geqslant 0} \frac{a_n}{n+1} z^{n+1} = \sum_{n\geqslant 1} \frac{a_{n-1}}{n} z^n .$$

- Une série entière, sa série dérivée et sa série primitive ont même rayon de convergence.
- **Dérivation d'une série entière de la variable réelle**

  Soit (a_n) une suite d'éléments de $\mathbb{C}$, et $\sum_{n\in\mathbb{N}} a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$. Posons, pour $x \in]-R\,;R[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$. Alors f est de classe $\mathscr{C}^1$ sur $]-R\,;R[$ et :

$$\forall\, x \in]-R\,;R[,\ f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n .$$

  Conséquence

  Avec les mêmes notations, la fonction f est de classe $\mathscr{C}^\infty$ sur sur $]-R\,;R[$ et, pour tout entier naturel k et pour tout $x \in]-R\,;R[$:

$$f^{(k)}(x) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} a_n x^{n-k} = \sum_{n=0}^{+\infty} \frac{(n+k)!}{n!} a_{n+k} x^n .$$

  Avec les mêmes notations on a aussi :

$$\forall\, k \in \mathbb{N},\ a_k = \frac{f^{(k)}(0)}{k!} .$$

- **Unicité du développement en série entière**

  Soient deux séries entières $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ et $g(x) = \sum_{n=0}^{+\infty} b_n x^n$, de rayons de convergence non nuls.

  Si f et g coïncident dans un voisinage de 0, on a $a_n = b_n$ pour tout entier n.

[S13.6] Intégration d'une série entière de la variable réelle

Soit (a_n) une suite d'éléments de $\mathbb{C}$, et $\sum_{n\in\mathbb{N}} a_n x^n$ une série entière de la variable réelle x, de rayon de convergence $R > 0$. Posons, pour $x \in]-R\,;R[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.

– Pour tout segment $[a\,;b] \subset]-R\,;R[$, on a : $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \sum_{n=0}^{+\infty} \int_a^b a_n x^n\,\mathrm{d}x$.

– En particulier, pour tout $x \in]-R\,;R[$, on a : $\displaystyle\int_0^x f(t)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n x^{n+1}}{n+1}$.

[S13.7] Fonction développable en série entière

On dit qu'une fonction $f\colon \mathbb{C} \to \mathbb{C}$ est développable en série entière dans un voisinage V de 0 s'il existe une série entière $\sum_{n\in\mathbb{N}} a_n z^n$ de rayon de convergence $R > 0$ telle que, pour tout $z \in V$, $f(z) = \sum_{n=0}^{+\infty} a_n z^n$.

D'après les résultats précédents, on a donc nécessairement, dans le cas d'une fonction d'une variable réelle :

- f est de classe $\mathscr{C}^\infty$ sur $]-R\,;R[$;
- $\forall\, n \in \mathbb{N},\ a_n = \dfrac{f^{(n)}(0)}{n!}$.

Ainsi, la série entière $\sum_{n\in\mathbb{N}} a_n x^n$ n'est autre que la série de Taylor de f en 0.

✓ Cette propriété n'admet pas de réciproque.

- Il existe des fonctions de classe $\mathscr{C}^\infty$ au voisinage de 0 dont la série de Taylor en 0 a un rayon de convergence nul.
- Il existe des fonctions de classe $\mathscr{C}^\infty$ au voisinage de 0 dont la série de Taylor en 0 a un rayon de convergence non nul, mais dont la somme ne coïncide pas avec f (sauf en 0).

[S13.8] Développements en série entière usuels

Les développements en série entière des fonctions usuelles sont rappelés page suivante. Quelques remarques concernant ce tableau.

- Les développements en série entière de $\sqrt{1+x}$ et de $\dfrac{1}{\sqrt{1+x}}$ n'y figurent pas, car ils s'obtiennent directement à partir du développement de $(1+x)^\alpha$ pour $\alpha = \pm\dfrac{1}{2}$.

  Cependant, il faut savoir écrire ces développements sous forme synthétique. On trouve :

$$\forall\, x \in]-1\,;1[,\ \sqrt{1+x} = 1 + \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{(2n-2)!}{2^{2n-1} n!(n-1)!} x^n\,,$$

  et

$$\forall\, x \in]-1\,;1[,\ \frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{+\infty} (-1)^n \frac{(2n)!}{2^{2n}(n!)^2} x^n.$$

- De même, le développement en série entière de Arcsin s'obtient simplement en intégrant celui de $\dfrac{1}{\sqrt{1-x^2}}$.

  Il faut savoir le retrouver ; on obtient :

$$\forall\, x]-1\,;1[,\ \operatorname{Arcsin}(x) = \sum_{n=0}^{+\infty} \frac{(2n)!}{2^{2n}(n!)^2} \frac{x^{2n+1}}{2n+1}\,.$$

Développements en série entière usuels

Fonction	Développement en série entière	Domaine
e^x	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{n!}$	$\mathbb{R}$
$\mathrm{sh}(x)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}$	$\mathbb{R}$
$\mathrm{ch}(x)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}$	$\mathbb{R}$
$\sin(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$\mathbb{R}$
$\cos(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!}$	$\mathbb{R}$
$\dfrac{1}{1-x}$	$\displaystyle\sum_{n=0}^{+\infty} x^n$	$]-1\,;1[$
$\dfrac{1}{1+x}$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n x^n$	$]-1\,;1[$
$\ln(1+x)$	$\displaystyle\sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^n}{n}$	$]-1\,;1]$
$\dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)$	$\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1}$	$]-1\,;1[$
$\mathrm{Arctan}(x)$	$\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$	$[-1\,;1]$
$(1+x)^\alpha$	$\displaystyle 1+\sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}\, x^n$	$]-1\,;1[$

Thème 14 - Intégration sur un intervalle quelconque

On considère dans ce chapitre des applications à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

[S14.1] Définitions

- Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale généralisée (ou impropre) $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si :

  $$\lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t \quad \text{existe.}$$

  S'il en est ainsi, cette intégrale est notée :

  $$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{[a;b[} f\,.$$

  Dans le cas contraire on dit que l'intégrale généralisée $\displaystyle\int_a^b f$ diverge.

- **Intégrale faussement impropre**

  Soit f continue par morceaux sur $[a\,;b[$ avec a et b *finis*, à valeurs dans $\mathbb{K}$.

  Si f admet une limite *finie* en b^-, alors l'intégrale $\displaystyle\int_{[a;b[} f$ converge.

  De plus, si l'on note $\tilde{f}$ le prolongement de f par continuité en b, on a :

  $$\int_{[a;b[} f = \int_a^b \tilde{f}\,.$$

  ✓ Attention à ne pas utiliser ce résultat dans le cas où $b = +\infty$!
  Une fonction définie sur un intervalle de la forme $[a\,;+\infty[$ peut avoir une limite en $+\infty$ sans que son intégrale converge. Il existe même des fonctions dont l'intégrale converge et qui n'ont pas de limite en $+\infty$.
  Ainsi, contrairement aux séries, il n'y a aucun lien entre limite en $+\infty$ et existence de l'intégrale dans le cas général.

- Soit f une fonction continue par morceaux sur $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$, et soit $c \in]a\,;b[$. L'intégrale généralisée $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si l'intégrale $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ converge et, dans ce cas :

  $$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t\,.$$

  ✓ Ainsi, pour une fonction continue par morceaux sur $[a\,;b[$, l'existence de l'intégrale est une notion locale au voisinage de b.

- Soit f une fonction continue par morceaux sur un intervalle $]a\,;b]$ avec $-\infty \leqslant a < b < +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale généralisée $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si :

  $$\lim_{x\to a^+} \int_x^b f(t)\,\mathrm{d}t \quad \text{existe.}$$

  S'il en est ainsi, cette intégrale est notée :

  $$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{]a;b]} f\,.$$

  Dans le cas contraire on dit que l'intégrale $\displaystyle\int_a^b f$ diverge.

- Soit f une application continue par morceaux sur un intervalle $]a\,;b[$ avec $-\infty \leqslant a < b \leqslant +\infty$ à valeurs dans $\mathbb{K}$. Les propriétés suivantes sont équivalentes :

  (i) il existe $c \in]a\,;b[$ tel que les intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ et $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ convergent ;

  (ii) pour tout $c \in]a\,;b[$, les intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ et $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ convergent.

  Si tel est le cas, la somme des deux intégrales généralisées $\displaystyle\int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$ ne dépend pas de c. On la note :

  $$\int_a^b f(t)\,\mathrm{d}t \quad \text{ou} \quad \int_a^b f \quad \text{ou} \quad \int_{]a;b[} f\,,$$

  et on dit alors que l'intégrale généralisée $\displaystyle\int_a^b f$ converge.

  On dira donc que l'intégrale $\displaystyle\int_a^b f$ diverge s'il existe $c \in]a\,;b[$ tel que *l'une au moins* des intégrales $\displaystyle\int_a^c f(t)\,\mathrm{d}t$ ou $\displaystyle\int_c^b f(t)\,\mathrm{d}t$ diverge.

[S14.2] Propriétés

Les propriétés qui suivent sont énoncées dans le cas d'une fonction continue par morceaux sur un intervalle de la forme $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$. On obtient bien sûr des résultats analogues pour les deux autres cas d'intégrale généralisée.

- Soit f et g deux fonctions continues par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{K}$ et λ, μ deux scalaires.

  On suppose que les intégrales $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ et $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ convergent.

  Alors l'intégrale généralisée $\displaystyle\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t$ converge et :

  $$\int_a^b (\lambda f + \mu g)(t)\,\mathrm{d}t = \lambda \int_a^b f(t)\,\mathrm{d}t + \mu \int_a^b g(t)\,\mathrm{d}t\,.$$

En d'autres termes : le sous-ensemble de $\mathscr{CM}([a\,;b[,\mathbb{K})$ formé des fonctions dont l'intégrale converge est un sous-espace vectoriel de $\mathscr{CM}([a\,;b[,\mathbb{K})$, et sur cet espace, l'application $f \mapsto \int_a^b f$ est une forme linéaire.

✓ Si f a une intégrale divergente sur $[a\,;b[$ et g une intégrale convergente, alors l'intégrale de $f+g$ est divergente.

On ne peut cependant rien dire a priori de la somme de deux intégrales divergentes.

- **Positivité**
  - Soit f une fonction continue par morceaux sur $[a\,;b[$, *à valeurs réelles.*

    Si $f \geqslant 0$ sur $[a\,;b[$ et si l'intégrale de f est convergente, alors $\int_a^b f(t)\,\mathrm{d}t \geqslant 0$.
  - Soient f et g deux fonctions continues par morceaux sur $[a\,;b[$, à valeurs réelles. Si $f(t) \leqslant g(t)$ pour tout $t \in [a\,;b[$ et si les intégrales de f et de g convergent, alors $\int_a^b f(t)\,\mathrm{d}t \leqslant \int_a^b g(t)\,\mathrm{d}t$.
  - Soit f une fonction *continue* sur $[a\,;b[$, *à valeurs réelles positives*, telle que l'intégrale de f sur $[a\,;b[$ est convergente.

    Alors : $\int_a^b f(t)\,\mathrm{d}t = 0 \implies f = 0$ sur $[a\,;b[$.
- **Cas d'une fonction à valeurs dans $\mathbb{C}$**

  Soit f une fonction continue par morceaux sur $[a\,;b[$, à valeurs dans $\mathbb{C}$.

  Pour que l'intégrale de f sur $[a\,;b[$ soit convergente, il faut et il suffit que les intégrales de $\mathcal{R}e\,(f)$ et de $\mathcal{I}m\,(f)$ le soient, et, dans ce cas on a

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^b \mathcal{R}e\,(f(t))\,\mathrm{d}t + \mathrm{i}\int_a^b \mathcal{I}m\,(f(t))\,\mathrm{d}t\,.$$

[S14.3] Règles de convergence pour les fonctions à valeurs positives

Les propriétés qui suivent sont énoncées dans le cas d'une fonction continue par morceaux sur un intervalle de la forme $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$. On obtient des résultats analogues pour une fonction continue par morceaux sur un intervalle de la forme $]a\,;b]$ avec $-\infty \leqslant a < b < +\infty$.

✓ Cependant, dans le cas d'une fonction définie sur un intervalle ouvert $]a\,;b[$, il faut faire deux études séparées, l'une au voisinage de a et l'autre au voisinage de b.

- Soit f continue par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles *positives.*

  Soit l'application $F\colon x \in [a\,;b[\longmapsto \int_a^x f(t)\,\mathrm{d}t\,.$

  Alors l'intégrale $\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si F est majorée et, dans ce cas, $\int_a^b f(t)\,\mathrm{d}t = \lim\limits_{x\to b^-} F(x)$.

  Dans le cas contraire, $\lim\limits_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t = +\infty\,.$

- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles.

  On suppose qu'il existe $c \in [a\,;b[$ tel que :

  $$\forall\, t \in [c\,;b[\,,\ 0 \leqslant f(t) \leqslant g(t)\,.$$

  – Si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.

  – Si $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ diverge, alors $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ diverge.

- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles, *positives* au voisinage de b.

  – Si $f \underset{b^-}{=} O(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.

  – Si $f \underset{b^-}{=} o(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge.

- Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles.

  On suppose g positive au voisinage de b. S'il existe une constante réelle $k \neq 0$ telle que $f \underset{b^-}{\sim} kg$, alors les intégrales de f et g sur $[a\,;b[$ sont de même nature.

[S14.4] Intégrales de référence

Pour pouvoir utiliser les critères de comparaison précédents, il faut connaître un certain nombre d'intégrales de référence.

- **Exponentielle et logarithme**

  . $\displaystyle\int_0^1 \ln t\,\mathrm{d}t$ converge (et $\displaystyle\int_0^1 \ln t\,\mathrm{d}t = -1$).

  . $\displaystyle\int_0^{+\infty} \mathrm{e}^{-at}\,\mathrm{d}t$ converge si et seulement si $a > 0$ (et dans ce cas $\displaystyle\int_0^{+\infty} \mathrm{e}^{-at}\,\mathrm{d}t = \frac{1}{a}$).

- **Fonctions de Riemann**

  Il s'agit des fonctions $t \mapsto \dfrac{1}{t^\alpha}$ pour $t > 0$ et $\alpha \in \mathbb{R}$.

  . $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ converge si et seulement si $\alpha > 1$.

  . $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t^\alpha}$ converge si et seulement si $\alpha < 1$.

  . Et plus généralement, si a et b sont des réels *finis* tels que $a < b$, alors :

  $$\int_a^b \frac{\mathrm{d}t}{(t-a)^\alpha} \quad \text{converge si et seulement si } \alpha < 1\,.$$

[S14.5] Changement de variable dans une intégrale généralisée

Soit f une fonction continue par morceaux sur un intervalle $]\alpha\,;\beta[$ à valeurs réelles ou complexes, et φ une bijection strictement monotone d'un intervalle $]a\,;b[$ sur $]\alpha\,;\beta[$, de classe $\mathscr{C}^1$ sur $]a\,;b[$.

Alors l'intégrale $\displaystyle\int_\alpha^\beta f$ est convergente si et seulement si l'intégrale $\displaystyle\int_a^b (f\circ\varphi)\varphi'$ l'est, et, dans ce cas :

$$\int_\alpha^\beta f(t)\,\mathrm{d}t = \int_a^b f\circ\varphi\,(u)\varphi'(u)\,\mathrm{d}u\,.$$

[S14.6] Intégration par parties

Soient f et g deux fonctions continues et de classe $\mathscr{C}^1$ par morceaux sur un intervalle $[a\,;b[$. Si $\lim\limits_{x\to b^-} f(t)g(t)$ existe , alors les intégrales $\displaystyle\int_a^b f'g$ et $\displaystyle\int_a^b fg'$ sont de même nature, et, lorsqu'elles convergent, on a :

$$\int_a^b f'(t)g(t)\,\mathrm{d}t = \lim_{t\to b^-} f(t)g(t) - f(a)g(a) - \int_a^b f(t)g'(t)\,\mathrm{d}t\,.$$

ce que l'on écrit encore $\displaystyle\int_a^b f'(t)g(t)\,\mathrm{d}t = \big[f(t)g(t)\big]_a^b - \int_a^b f(t)g'(t)\,\mathrm{d}t.$

Ce théorème s'adapte sans difficulté au cas d'un intervalle de la forme $]a\,;b]$ (il faut alors que $\lim\limits_{t\to a^+} f(t)g(t)$ existe) ou au cas d'un intervalle de la forme $]a\,;b[$ (et il faut alors que les *deux* limites $\lim\limits_{t\to b^-} f(t)g(t)$ et $\lim\limits_{t\to a^+} f(t)g(t)$ existent).

> ✓ Pour appliquer ce théorème il est *indispensable* de vérifier proprement l'existence de la limite. En cas de doute, on reviendra à la formule sur un segment, et à la fin des calculs on fera un (ou des) passage(s) à la limite.

[S14.7] Intégrales absolument convergentes

Lorsqu'une fonction n'est pas à valeurs réelles positives (ou de signe constant), les théorèmes de comparaison vus plus haut ne s'appliquent pas. On peut cependant s'y ramener dans certains cas grâce à la définition et au théorème suivants.

- **Définition**

  Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  On dit que l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente si l'intégrale $\displaystyle\int_a^b |f(t)|\,\mathrm{d}t$ est convergente ($|f(t)|$ désigne la valeur absolue de $f(t)$ si f est à valeurs réelles, et son module si f est à valeurs dans $\mathbb{C}$).

  Lorsque l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente, on dit aussi que f est intégrable sur $[a\,;b[$.

  On a bien sûr une définition analogue dans le cas d'une fonction définie sur un intervalle de la forme $]a\,;b]$ ou $]a\,;b[$.

- **Théorème**

  Soit f une fonction continue par morceaux sur un intervalle $[a\,;b[$ avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{K}$.

  Si l'intégrale $\displaystyle\int_a^b f$ est absolument convergente, elle est convergente, et alors :

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(t)|\,\mathrm{d}t\,.$$

  ✓ La réciproque du théorème précédent est *fausse*.

  Il existe en effet des intégrales qui sont convergentes sans être absolument convergentes (une telle intégrale est dite semi-convergente). Exemple : $\displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t$ est convergente mais pas absolument convergente.

- Le théorème précédent permet d'utiliser certains critères de comparaison même lorsqu'on ne connaît pas le signe de la fonction ; on a en particulier le résultat suivant.

  Soient f et g continues par morceaux sur $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$.

  On suppose (seulement) g à valeurs réelles *positives* au voisinage de b.

  Si $f \underset{b^-}{=} O(g)$ ou si $f \underset{b^-}{=} o(g)$ et si $\displaystyle\int_a^b g(t)\,\mathrm{d}t$ converge, alors $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ est absolument convergente.

[S14.8] Intégration des relations de comparaison

Les théorèmes suivants permettent d'obtenir une estimation du « reste » d'une intégrale convergente ou des « intégrales partielles » d'une intégrale divergente.

Ces théorèmes sont à mettre en parallèle avec ceux concernant la sommation des relations de comparaison pour les séries, rappelés en **[S11.12]**.

Dans ce qui suit , $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et $[a\,;b[$ un intervalle de $\mathbb{R}$ avec $a < b \leqslant +\infty$. On obtient des résultats similaires dans le cas de fonctions définies sur un intervalle $]a\,;b]$ avec $-\infty \leqslant a < b$.

- **Cas des fonctions intégrables**

  . Soit f une fonction continue par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{K}$ et g une fonction continue par morceaux sur $[a\,;b[$ à valeurs *réelles positives*.

  On suppose que g est intégrable sur $[a\,;b[$ et que $f(x) \underset{x\to b^-}{=} O\big(g(x)\big)$ $\big($respectivement $f(x) \underset{x\to b^-}{=} o\big(g(x)\big)\big)$.

  Alors f est intégrable sur $[a\,;b[$ et

$$\int_x^b f(t)\,\mathrm{d}t \underset{x\to b^-}{=} O\left(\int_x^b g(t)\,\mathrm{d}t\right)$$

$$\left(\text{respectivement } \int_x^b f(t)\,\mathrm{d}t \underset{x\to b^-}{=} o\left(\int_x^b g(t)\,\mathrm{d}t\right)\right).$$

- Soit f une fonction continue par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{R}$ et g une fonction continue par morceaux sur $[a\,;b[$ à valeurs *réelles positives.*
  On suppose que g est intégrable sur $[a\,;b[$ et que $f(x) \underset{x\to b^-}{\sim} g(x)$.

  Alors f est intégrable sur $[a\,;b[$ et l'on a : $\displaystyle\int_x^b f(t)\,\mathrm{d}t \underset{x\to b^-}{\sim} \int_x^b g(t)\,\mathrm{d}t$.

- **Cas des fonctions non intégrables**
  - Soit f une fonction continue par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{K}$ et g une fonction continue par morceaux sur $[a\,;b[$ à valeurs *réelles positives.*
    On suppose que : g n'est *pas* intégrable sur $[a\,;b[$ et que $f(x) \underset{x\to b^-}{=} O\big(g(x)\big)$ (respectivement $f(x) \underset{x\to b^-}{=} o\big(g(x)\big)$).

    Alors :
    $$\int_a^x f(t)\,\mathrm{d}t \underset{x\to b^-}{=} O\left(\int_a^x g(t)\,\mathrm{d}t\right)$$
    $$\left(\text{respectivement } \int_a^x f(t)\,\mathrm{d}t \underset{x\to b^-}{=} o\left(\int_a^x g(t)\,\mathrm{d}t\right)\right).$$
  - Soit f une fonction continue par morceaux sur $[a\,;b[$ à valeurs dans $\mathbb{R}$ et g une fonction continue par morceaux sur $[a\,;b[$ à valeurs *réelles positives.*
    On suppose que g n'est *pas* intégrable sur $[a\,;b[$ et que $f(x) \underset{x\to b^-}{\sim} g(x)$.

    Alors f n'est pas intégrable sur $[a\,;b[$ et $\displaystyle\int_a^x f(t)\,\mathrm{d}t \underset{x\to b^-}{\sim} \int_a^x g(t)\,\mathrm{d}t$.

[S14.9] Suite de fonctions intégrables

Contrairement à l'intégration sur un segment (cf. **[S12.2]**), même la convergence uniforme d'une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ vers une fonction f ne suffit pas, lorsque l'intervalle I est quelconque, à assurer que :

$$\lim_{n\to+\infty} \int_I f_n = \int_I f\,.$$

Il nous faut un résultat plus puissant ; il s'agit du **théorème de convergence dominée de Lebesgue.**

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{K}$. On suppose que :

– la suite de fonctions f_n converge simplement sur I vers une fonction f continue par morceaux sur I ;

– il existe une fonction φ continue par morceaux sur I, à valeurs réelles positives, et intégrable sur I, telle que : $\forall\, n \in \mathbb{N}$, $|f_n| \leqslant \varphi$ sur I (*hypothèse de domination*).

Alors les f_n sont intégrables sur I, f est intégrable sur I, et

$$\int_I f \;=\; \lim_{n\to+\infty} \int_I f_n$$

(*l'existence de cette limite est assurée par le théorème*).

[S14.10] Série de fonctions intégrables

Deux théorèmes sont disponibles pour calculer l'intégrale de la somme d'une série de fonctions, en plus de celui déjà vu en **[S12.10]** dans le cas où il y a convergence uniforme sur un segment.

- **Théorème d'intégration terme à terme**
  Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{C}$. On suppose que :
  - $\forall\, n \in \mathbb{N}$, u_n est intégrable sur I ;
  - la série de fonctions $\sum u_n$ converge simplement sur I, et sa fonction somme s est continue par morceaux sur I ;
  - la série $\sum \int_I |u_n|$ converge.

  Alors $s = \sum_{n=0}^{+\infty} u_n$ est intégrable sur I, et $\int_I s = \sum_{n=0}^{+\infty} \int_I u_n$ (cette série étant absolument convergente).
- **Application du théorème de convergence dominée**
  Soit $(u_n)_{n\in\mathbb{N}}$ une suite de fonctions continues par morceaux sur un intervalle I, à valeurs dans $\mathbb{C}$. On notera, pour $n \in \mathbb{N}$, $s_n = \sum_{k=0}^{n} u_k$. On suppose que :
  - la série de fonctions $\sum u_n$ converge simplement sur I, et sa fonction somme $s = \sum_{n=0}^{+\infty} u_n$ est continue par morceaux sur I ;
  - il existe une fonction φ continue par morceaux sur I, à valeurs réelles positives, et intégrable sur I, telle que : $\forall\, n \in \mathbb{N}$, $|s_n| \leqslant \varphi$ sur I (*hypothèse de domination*).

  Alors toutes les fonctions u_n, ainsi que s, sont intégrables sur I, et $\int_I s = \sum_{n=0}^{+\infty} \int_I u_n$ (*la convergence de cette série étant assurée par le théorème*).

[S14.11] Intégrales dépendant d'un paramètre

- **Continuité**
  Soit f une fonction définie sur $A \times I$, où A est une partie d'un espace vectoriel normé de dimension finie et où I est un intervalle de $\mathbb{R}$, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que :
  - $\forall\, x \in A$, $t \mapsto f(x,t)$ est continue par morceaux sur I ;
  - $\forall\, t \in I$, la fonction $x \mapsto f(x,t)$ est continue sur A ;
  - il existe une fonction φ, continue par morceaux sur I, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* I telle que :

$$\forall\,(x,t) \in A \times I,\ \big|f(x,t)\big| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

Alors la fonction $g\colon x \mapsto \int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur A.

Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in V \times I$, pour tout voisinage V de tout point de A (domination locale).

> ✓ La remarque précédente est très importante, et simplifie souvent beaucoup la recherche d'une fonction dominante.

- **Passage à la limite**

Le résultat suivant étend le théorème de convergence dominée.

Soit f une fonction définie sur $A \times I$, où A et I sont des intervalles de $\mathbb{R}$, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $a \in \overline{A}$. On suppose que :

- $\forall x \in A$, $t \mapsto f(x,t)$ est continue par morceaux sur I ;
- $\forall t \in I$, $\lim\limits_{x \to a} f(x,t) = \ell(t)$, avec ℓ continue par morceaux sur I ;
- il existe une fonction φ, continue par morceaux sur I, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* I telle que :

$$\forall (x,t) \in A \times I,\ \big|f(x,t)\big| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

Alors ℓ est intégrable sur I et : $\lim\limits_{x \to a} \int_I f(x,t)\,\mathrm{d}t = \int_I \ell(t)\,\mathrm{d}t$.

Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in V \times I$, avec voisinage V de a dans A.

- **Dérivabilité**

Soit f une fonction définie sur $J \times I$, où I et J sont des intervalles réels, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que :

- $\forall x \in J$, $t \mapsto f(x,t)$ est continue par morceaux et intégrable sur I ;
- $\forall t \in I$, la fonction $x \mapsto f(x,t)$ est dérivable sur J ;
- pour tout $x \in J$, la fonction $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur I ;
- pour tout $t \in I$, la fonction $x \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur J ;
- il existe une fonction φ, continue par morceaux sur I, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* I telle que :

$$\forall (x,t) \in J \times I,\ \left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t)\ (\textit{hypothèse de domination}).$$

Alors la fonction $g\colon x \mapsto \int_I f(x,t)\,\mathrm{d}t$ est de classe $\mathscr{C}^1$ sur J, et :

$$\forall x \in J,\ g'(x) = \int_I \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t \quad (\textit{formule de Leibniz}).$$

Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in K \times I$, pour tout segment K inclus dans J.

> ✓ Si l'on vous demande de prouver que la fonction g est de classe $\mathscr{C}^1$, appliquez directement ce dernier théorème sans passer par le premier !

- **Généralisation**

  Le résultat suivant se démontre aisément par récurrence à partir du précédent.

  Soit f une fonction définie sur $J \times I$, où I et J sont des intervalles réels, à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $n \in \mathbb{N}^*$.

  On suppose que :

  – $f, \dfrac{\partial f}{\partial x}, \ldots, \dfrac{\partial^n f}{\partial x^n}$ existent, sont continues par rapport à x sur J et continues par morceaux par rapport à t sur I ;

  – $\forall x \in J$, les fonctions $t \mapsto f(x,t)$, $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$, $\ldots$, $t \mapsto \dfrac{\partial^{n-1} f}{\partial x^{n-1}}(x,t)$ sont intégrables sur I ;

  – il existe une fonction φ, continue par morceaux sur I, à valeurs dans $\mathbb{R}_+$ et *intégrable sur* I telle que :

$$\forall (x,t) \in J \times I, \ \left|\frac{\partial^n f}{\partial x^n}(x,t)\right| \leqslant \varphi(t) \ (\textit{hypothèse de domination}).$$

  Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est de classe $\mathscr{C}^n$ sur J, et :

$$\forall k \in [\![0\,;n]\!], \forall x \in J, \ g^{(k)}(x) = \int_I \frac{\partial^k f}{\partial x^k}(x,t)\,\mathrm{d}t\,.$$

  Remarque : le résultat de ce théorème reste vrai si l'on suppose seulement que l'hypothèse de domination est vérifiée pour $(x,t) \in K \times I$, pour tout segment K de J.

  ✓ Si l'on vous demande de montrer que g est de classe $\mathscr{C}^n$ voire $\mathscr{C}^\infty$, il faut utiliser directement ce dernier théorème.

Thème 15 - Espaces probabilisés

Avant d'aborder les probabilités, il nous a paru indispensable de faire quelques rappels de 1re année concernant les dénombrements et les coefficients binomiaux. Ce chapitre utilise aussi les résultats sur les ensembles dénombrables, les familles sommables et les séries doubles qui ont été rappelés en **[S11.13]** et **[S11.14]**.

[S15.1] Dénombrement

- **p-listes**
  Soient n et p deux entiers strictement positifs et E un ensemble à n éléments.
  Une p-liste d'éléments de E est un p-uplet $(x_1, \ldots, x_p)$ d'éléments de E.

  ✓ Dans une p-liste, les éléments ne sont pas nécessairement distincts, mais le rang des éléments dans la p-liste est important.

  Le nombre de p-listes d'un ensemble E à n éléments est $\mathrm{Card}(E^p) = n^p$.

- **p-listes d'éléments distincts de E, ou arrangements**
  Soient n et p deux entiers tels que $0 < p \leqslant n$ et E un ensemble à n éléments.
  Une p-liste d'éléments distincts de E ou p-arrangement est une p-liste $(x_1, \ldots, x_p)$ d'éléments de E telle que $x_i \neq x_j$ pour $i \neq j$.

  ✓ Dans un p-arrangement, les éléments sont tous distincts, et le rang des éléments dans la liste est important.

  Si $\mathrm{Card}\, E = n$, le nombre de p-listes d'éléments distincts de E est le nombre noté A_n^p égal à :

$$A_n^p = n(n-1)\cdots(n-p+1) = \frac{n!}{(n-p)!}\cdot$$

- **Permutations**
  Une permutation d'un ensemble E de cardinal n est une n-liste d'éléments distincts de E.
  Le nombre de permutations d'un ensemble E de cardinal n est : $A_n^n = n!$.

- **Parties d'un ensemble**
  Soit E un ensemble de cardinal n ($n \in \mathbb{N}$) et $\mathcal{P}(E)$ l'ensemble des parties de E.
  Alors : $\mathrm{Card}\, \mathcal{P}(E) = 2^n$.

- **Parties de E à p éléments, combinaisons**
  Soit $p \in \mathbb{N}$. On appelle combinaison à p éléments d'un ensemble E toute partie de E contenant p éléments.

  ✓ Dans une combinaison, les éléments sont tous distincts et leur rang est sans importance.

  Si $\mathrm{Card}\, E = n$, le nombre de combinaisons à p éléments de E ($p \leqslant n$) se note $\binom{n}{p}$ et l'on a :

$$\binom{n}{p} = \frac{1}{p!} A_n^p = \frac{n!}{p!(n-p)!}\cdot$$

On posera $\binom{0}{0} = 1$.

Si $n \in \mathbb{N}$ et $p \in \mathbb{Z}$ sont tels que $p < 0$ ou $p > n$, on posera $\binom{n}{p} = 0$.

- **Propriétés des coefficients binomiaux**

1. Pour $n \in \mathbb{N}$, $\binom{n}{0} = \binom{n}{n} = 1$; $\binom{n}{1} = n$; $\binom{n}{2} = \dfrac{n(n-1)}{2}$.

2. Pour $n \in \mathbb{N}$ et $k \in \mathbb{Z}$, $\binom{n}{k} = \binom{n}{n-k}$.

3. Formule sans nom :

$$\forall n \in \mathbb{N},\ \forall k \in \mathbb{Z},\ k\binom{n}{k} = n\binom{n-1}{k-1}.$$

4. Formule de Pascal :

$$\forall n \in \mathbb{N},\ \forall k \in \mathbb{Z},\ \binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

5. Généralisation de la formule de Pascal :

$$\text{pour } 0 \leqslant p \leqslant n,\ \sum_{k=p}^{n} \binom{k}{p} = \binom{n+1}{p+1}.$$

6. Formule de Vandermonde :

$$\forall (n,m,k) \in \mathbb{N}^3,\ \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}.$$

7. Cas particulier : $\displaystyle\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}$.

8. Formule du binôme :

$$\forall n \in \mathbb{N},\ \forall (a,b) \in \mathbb{C}^2,\quad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

- **Applications de E_p dans F_n**

Soient p et n deux entiers strictement positifs, E_p un ensemble de cardinal p dont les éléments sont notés $(e_1, \ldots, e_p)$ et F_n un ensemble de cardinal n.

À toute p-liste $(y_1, \ldots, y_p)$ d'éléments de F_n correspond une et une seule application de E_p dans F_n définie par : $f(e_i) = y_i$ pour tout i.

On en déduit les résultats suivants.

- Le nombre d'applications de E_p dans F_n est égal à : n^p.
- Le nombre d'injections de E_p dans F_n $(n \geqslant p)$ est égal à : A_n^p.
- Le nombre de bijections de E_n dans F_n $(n = p)$ est égal à : $A_n^n = n!$.

[S15.2] Espace probabilisable

- Soient Ω un ensemble appelé univers et $\mathcal{A}$ une partie de $\mathcal{P}(\Omega)$, c'est-à-dire un ensemble de parties de Ω.
  On dit que $\mathcal{A}$ est une tribu ou une σ-algèbre de parties de Ω si :
  - $\Omega \in \mathcal{A}$;
  - si $A \in \mathcal{A}$ alors $\overline{A} \in \mathcal{A}$ ($\overline{A}$ désigne le complémentaire de A dans Ω) ;
  - pour toute partie I de $\mathbb{N}$, et toute famille $(A_i)_{i\in I}$ d'éléments de $\mathcal{A}$, $\bigcup\limits_{i\in I} A_i \in \mathcal{A}$.

  Les éléments de $\mathcal{A}$ sont appelés des évènements, et le couple $(\Omega,\mathcal{A})$ s'appelle un espace probabilisable.
- **Propriétés**
  Soient Ω un ensemble et $\mathcal{A}$ une tribu de partie de Ω.
  - $\emptyset \in \mathcal{A}$.
  - Si A et B appartiennent à $\mathcal{A}$, alors $A \cup B$, $A \cap B$ et $A \setminus B$ sont dans $\mathcal{A}$.
  - Si I est une partie de $\mathbb{N}$ et si pour tout $i \in I$, $A_i \in \mathcal{A}$ alors $\bigcap\limits_{i\in I} A_i \in \mathcal{A}$.
- **Un peu de vocabulaire**

Terminologie probabiliste	Terminologie ensembliste
évènement certain	Ω
évènement impossible	$\emptyset$
évènement élémentaire ω	singleton $\{\omega\}$
évènement contraire de A : $\bar{A}$	complémentaire de A dans Ω
évènement A ou B	$A \cup B$
évènement A et B	$A \cap B$
les évènements A et B sont incompatibles	$A \cap B = \emptyset$
A implique B	$A \subset B$
système complet d'évènements	partition

[S15.3] Espace probabilisé

- Soit $(\Omega,\mathcal{A})$ un espace probabilisable. On appelle probabilité sur $(\Omega,\mathcal{A})$ toute application P de $\mathcal{A}$ dans $[0\,;1]$ vérifiant :
  (i) $P(\Omega) = 1$;
  (ii) si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements deux à deux incompatibles alors :
  $$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \qquad (\sigma\text{-additivité})$$
  (cette écriture sous-entend que la série converge).

Le triplet $(\Omega,\mathcal{A},P)$ est alors appelé espace probabilisé.

- **Germe de probabilité**
  Soit Ω un ensemble au plus dénombrable, et $\{p_\omega \mid \omega \in \Omega\}$ une famille sommable de réels positifs telle que $\sum_{\omega\in\Omega} p_\omega = 1$.
  Alors il existe une unique probabilité P sur $\big(\Omega,\mathcal{P}(\Omega)\big)$ telle que :
  $$\forall\, \omega \in \Omega,\ P(\{\omega\}) = p_\omega\,.$$
  Cette probabilité est alors définie par :
  $$\forall\, A \in \mathcal{P}(\Omega),\ P(A) = \sum_{\omega\in A} p_\omega\,.$$

  ✓ Ce théorème permet donc de généraliser au cas où Ω est dénombrable ce que vous avez vu en 1re année dans le cas fini : pour définir une probabilité, il suffit de la définir pour les évènements élémentaires, sous réserve bien sûr que toutes ces probabilités soient des réels positifs et que leur somme soit une famille sommable de somme 1.

- **Propriétés**
  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et soient A et B deux évènements.
  - $P(\emptyset) = 0$.
  - $P(\overline{A}) = 1 - P(A)$.
  - Si A et B sont incompatibles, $P(A\cup B) = P(A) + P(B)$.
  - Plus généralement : $P(A\cup B) = P(A) + P(B) - P(A\cap B)$.
  - $P(A\setminus B) = P(A) - P(A\cap B)$.
  - Si $A\subset B$ alors $P(A)\leqslant P(B)$ (*croissance*).
  - Si $(A_i)_{i\in I}$ (I partie de $\mathbb{N}$) est un système complet d'évènements, $\sum_{i\in I} P(A_i) = 1$.

- Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé.
  - Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'évènements de $\mathcal{A}$ ($A_n\subset A_{n+1}$) alors la suite $\big(P(A_n)\big)_{n\in\mathbb{N}}$ converge et :
    $$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) \quad (\textit{continuité croissante}).$$
  - Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'évènements de $\mathcal{A}$ ($A_{n+1}\subset A_n$) alors la suite $\big(P(A_n)\big)_{n\in\mathbb{N}}$ converge et :
    $$P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) \quad (\textit{continuité décroissante}).$$

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements de $\mathcal{A}$ alors :
  $$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right) \quad,\quad P\left(\bigcap_{k=0}^{+\infty} A_k\right) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right).$$

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'évènements de $\mathcal{A}$ alors :
  - Pour tout $n \in \mathbb{N}$, $P\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{n=0}^{n} P(A_n)$ (*sous-additivité finie*).
  - $P\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{n=0}^{+\infty} P(A_n)$ (*sous-additivité*), cette dernière somme pouvant être $+\infty$.
- Un évènement $A \in \mathcal{A}$ est dit quasi-impossible si $P(A) = 0$ et quasi-certain si $P(A) = 1$.

  Un système quasi-complet d'évènements est une famille $(A_i)_{i\in I}$ ($I \subset \mathbb{N}$) d'éléments de $\mathcal{A}$, non vides, deux à deux incompatibles, telle que $P\left(\bigcup_{i\in I} A_i\right) = 1$.

[S15.4] Probabilité conditionnelle

- Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et B un évènement de probabilité non nulle.

  Pour tout évènement A, on pose $P_B(A) = \dfrac{P(A \cap B)}{P(B)}$.

  Alors P_B est une probabilité sur $(\Omega,\mathcal{A})$ appelée probabilité conditionnelle relative à B ou encore probabilité sachant B. $P_B(A)$ est aussi noté $P(A \mid B)$.

  Si $P(A) \neq 0$, on peut aussi écrire $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$ et donc :

$$P(A) \times P(B \mid A) = P(B) \times P(A \mid B).$$

  Comme P_B est une probabilité, toutes les propriétés vues en **[S15.3]** peuvent lui être appliquées (par exemple : $P\left(\overline{A} \mid B\right) = 1 - P(A \mid B)$).
- **Formule des probabilités composées**

  Directement d'après la définition, on a pour A, $B \in \mathcal{A}$:

$$P(A \cap B) = P(B)P(A \mid B).$$

  Plus généralement, si $(A_i)_{1\leqslant i\leqslant n}$ est une famille d'évènements tels que $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \neq 0$, on a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)\ldots P_{A_1\cap A_2\cap\ldots\cap A_{n-1}}(A_n),$$

  ce qui s'écrit aussi :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)\,P(A_3|A_1 \cap A_2)\ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$
- **Formule des probabilités totales**

  Soit $(A_i)_{i\in I}$ un système quasi-complet d'évènements de probabilités non nulles. Alors pour tout évènement B on a :

$$P(B) = \sum_{i\in I} P(A_i \cap B) = \sum_{i\in I} P(B \mid A_i)\,P(A_i)$$

(lorsque I est infini dénombrable, cela sous-entend que la famille écrite est sommable).

On peut interpréter cette formule en disant que B est un effet provenant de différentes causes A_i.

Un cas particulier important de la formule des probabilités totales est celui où l'on prend comme système complet d'évènements un système de la forme $\{A,\overline{A}\}$. La formule s'écrit alors :

$$P(B) = P(B \mid A)\,P(A) + P\left(B \mid \overline{A}\right)P(\overline{A})\,.$$

- **Formule de Bayes**

  On a vu que, si A et B sont deux évènements de probabilités non nulles, on a :

  $$P(A \mid B) = \frac{P(A) \times P(B \mid A)}{P(B)} \qquad (*)$$

  Si maintenant on considère un système complet ou quasi-complet d'évènements $(A_i)_{i\in I}$ (avec toujours I partie de $\mathbb{N}$), et si A est l'un de ces évènements, soit $A = A_j$ pour un $j \in I$, la formule des probabilités totales appliquée à l'évènement B donne la **formule de Bayes**.

  > Soient $(A_i)_{i\in I}$ un système quasi-complet d'évènements de probabilités non nulles et B un évènement de probabilité non nulle. On a :
  >
  > $$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\displaystyle\sum_{i\in I} P(A_i)P(B \mid A_i)}\,.$$

  Cette formule s'appelle aussi formule de probabilités des causes.

  > ✓ Cette formule est affreuse ! Il vaut mieux retenir la formule $(*)$, beaucoup plus simple et qui est une conséquence directe de la définition d'une probabilité conditionnelle, puis obtenir la valeur de $P(B)$ par la formule des probabilités totales.

[S15.5] Indépendance d'évènements

- Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé. On dit que deux évènements A et B sont indépendants pour la probabilité P si :

  $$P(A \cap B) = P(A) \times P(B)\,.$$

  On a donc alors $P(B \mid A) = P(B)$ et $P(A \mid B) = P(A)$.

  Si A et B sont deux évènements indépendants alors $\overline{A}$ et B, A et $\overline{B}$ et $\overline{A}$ et $\overline{B}$ le sont aussi.

  > ✓ Ne pas confondre « évènements indépendants » et « évènements incompatibles » ! D'ailleurs, il est facile de vérifier que deux évènements incompatibles sont indépendants si et seulement si l'un d'entre eux est quasi-impossible.
  >
  > De plus, la notion d'indépendance dépend de la probabilité P alors que la notion d'incompatibilité est purement ensembliste.

- Soient $(A_i)_{i\in I}$ une famille d'évènements (avec $I \subset \mathbb{N}$, fini ou non).
  - On dit que les évènements sont deux à deux indépendants pour la probabilité P si pour tout $i \neq j$ A_i et A_j sont indépendants.
  - On dit que les évènements sont mutuellement indépendants (ou tout simplement indépendants) pour la probabilité P si pour *tout* ensemble fini d'indices $J \subset I$, $P\left(\bigcap_{j\in J}\right) A_j = \prod_{j\in J} P(A_j)$.

    ✓ Il est clair que si des évènements sont mutuellement indépendants, ils le sont deux à deux, mais la réciproque est fausse.

- Soient $(A_i)_{i\in I}$ une famille d'évènements indépendants (avec $I \subset \mathbb{N}$, fini ou non). Soit $(B_i)_{i\in I}$ une famille d'évènements telle que :

$$\forall\, i \in I,\ B_i = A_i \quad \text{ou} \quad B_i = \overline{A_i}\,.$$

  Alors les B_i sont indépendants.

Thème 16 - Variables aléatoires discrètes

[S16.1] Variables aléatoires discrètes

- **Définitions**
  - Soit $(\Omega,\mathcal{A})$, un espace probabilisable. On appelle variable aléatoire discrète définie sur $(\Omega,\mathcal{A})$ à valeurs dans un ensemble E toute application X de Ω dans E telle que :
    - son image $X(\Omega)$ est au plus dénombrable (on écrira donc $X(\Omega) = \{x_i, i \in I\}$ avec $I \subset \mathbb{N}$) ;
    - pour tout $x \in X(\Omega)$, l'ensemble $X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}$ appartient à $\mathcal{A}$ (c'est-à-dire est un évènement).
      Cet ensemble se note, en abrégé : $(X = x)$.

    On a alors, plus généralement, le résultat suivant :
    pour toute partie $A \subset E$, l'ensemble $X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}$ est un évènement de $\mathcal{A}$, que l'on note en abrégé : $(X \in A)$.
  - Si $X(\Omega)$ est un ensemble fini, on dit que X est une variable aléatoire finie.
  - Si $X(\Omega)$ est une partie de $\mathbb{R}$, on dit que X est une variable aléatoire réelle.
  - Si X est constante, on dit que c'est une variable aléatoire certaine.
- **Fonction d'une variable aléatoire**
  Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$ à valeurs dans un ensemble E et f une fonction définie sur $X(\Omega)$ à valeurs dans un ensemble F.
  Alors $f \circ X$ est une variable aléatoire discrète sur $(\Omega,\mathcal{A},P)$ (on la notera simplement $f(X)$).
- **Loi d'une variable aléatoire**
  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé, et X une variable aléatoire discrète à valeurs dans un ensemble E.
  L'application $f : \begin{cases} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P(X = x) \end{cases}$ s'appelle la loi de probabilité de X.

  ✓ Pour répondre à la question « déterminer la loi de X », il faut commencer par donner clairement $X(\Omega)$. Ensuite, pour tout élément x_i de cet ensemble $X(\Omega)$, il faut donner $P(X = x_i)$.

  Le résultat suivant est extrêmement important et permet en particulier de vérifier la cohérence des résultats obtenus.
  Soit X une variable aléatoire discrète. Si $X(\Omega) = \{x_i,\ i \in I\}$ alors la famille d'évènements $(X = x_i)_{i\in I}$ est un système complet d'évènements. En particulier on a $\sum_{i\in I} P(X = x_i) = 1$.
  De plus, puisque $(X = x_i)_{i\in I}$ est un système complet d'évènements, on peut appliquer la formule des probabilités totales pour n'importe quel évènement A :

$$P(A) = \sum_{i\in I} P(A \cap (X = x_i)) = \sum_{i\in I} P(A \mid X = x_i)\, P(X = x_i)\,.$$

- **Germe d'une variable aléatoire**
  Soit $(\Omega,\mathcal{A})$ un espace probabilisable, $\{x_i,\ i\in I\}$ un ensemble au plus dénombrable, et $(p_i)_{i\in I}$ une famille sommable de réels positifs telle que $\sum_{i\in I} p_i = 1$.
  Alors il existe une probabilité P sur $(\Omega,\mathcal{A})$ et une variable aléatoire discrète X sur $(\Omega,\mathcal{A},P)$, à valeurs dans $\{x_i,\ i\in I\}$, telles que :
  $$\forall\, i\in I,\ \ P(X=x_i)=p_i\,.$$
  ✓ Ce résultat permet ainsi d'étudier une variable aléatoire définie par sa loi, sans avoir à préciser l'espace probabilisé sur lequel on travaille.

- **Fonction de répartition d'une variable aléatoire réelle**
  Soit X une variable aléatoire discrète réelle. On appelle fonction de répartition de X l'application $F\colon \mathbb{R}\to\mathbb{R}$ définie par :
  $$F(x)=P(X\leqslant x)\,.$$
  Propriétés
  - F est une fonction en escalier ;
  - $\forall\, x\in\mathbb{R},\ F(x)\in[0\,;1]$;
  - F est croissante ;
  - $\lim\limits_{x\to-\infty} F(x)=0$ et $\lim\limits_{x\to+\infty} F(x)=1$;
  - F est continue à droite en tout point ;
  - si a et b sont des réels tels que $a<b$, $P(a<X\leqslant b)=F(b)-F(a)$;
  - en notant $X(\Omega)=\{x_i,\ i\in I\}$, les x_i étant rangés par ordre croissant, pour tout $i\in I$ tel que $i-1\in I$ (on a donc $x_{i-1}<x_i$), on a :
    $$P(X=x_i)=F(x_i)-F(x_{i-1})\,.$$

[S16.2] Lois discrètes finies usuelles

- **Loi uniforme**
  On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi uniforme sur $[\![1\,;n]\!]$, et l'on note $X\hookrightarrow\mathcal{U}([\![1\,;n]\!])$ si :
  $$\text{(i) } X(\Omega)=[\![1\,;n]\!] \qquad \text{(ii) } \forall\, x\in X(\Omega),\ P(X=x)=\frac{1}{n}\cdot$$
  Situation : tirage au hasard d'un entier parmi $[\![1\,;n]\!]$ (ou une boule dans une urne…), tous les tirages étant équiprobables.

- **Loi de Bernoulli**
  Soit $p\in[0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi de Bernoulli de paramètre p, et l'on note $X\hookrightarrow\mathcal{B}(p)$ si :
  $$\text{(i) } X(\Omega)=\{0,1\} \qquad \text{(ii) } P(X=1)=p \ \text{ et } \ P(X=0)=1-p\,.$$
  Les cas $p=0$ et $p=1$ n'étant pas intéressants, on considère en général $p\in\,]0\,;1[$.
  Situation : expérience de type succès-échec (pile ou face, tirage d'une boule noire ou blanche…).

- **Loi binomiale**

Soit n un entier non nul et p un réel dans $[0\,;1]$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi binomiale de paramètres n et p, et l'on note $X \hookrightarrow \mathcal{B}(n,p)$ si :

$$\text{(i) } X(\Omega) = [\![0\,;n]\!] \qquad \text{(ii) } \forall k \in [\![0\,;n]\!],\ P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}\,.$$

On note traditionnellement $q = 1-p$ de sorte que $P(X=k) = \binom{n}{k} p^k q^{n-k}$.

Situation : expérience de type succès-échec que l'on effectue n fois dans les mêmes conditions ; X désigne le nombre de fois où l'expérience a amené un succès (lancers de pièces, tirage avec remise de boules dans une urne...).

[S16.3] Lois discrètes infinies usuelles

- **Loi géométrique**

Soit $p \in\,]0\,;1[$. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi géométrique de paramètre p, et l'on note $X \hookrightarrow \mathcal{G}(p)$ si :

$$\text{(i) } X(\Omega) = \mathbb{N}^* \qquad \text{(ii) } \forall n \in \mathbb{N}^*,\ P(X=n) = (1-p)^{n-1} p\,.$$

Il est facile de vérifier que $\displaystyle\sum_{n=1}^{+\infty} P(X=n)$ est bien égal à 1 (série géométrique).

Situation : expérience de type succès-échec que l'on répète dans des conditions identiques ; X désigne le nombre d'épreuves effectuées jusqu'à ce que l'on obtienne un succès pour la première fois.

Une propriété importante de la loi géométrique est d'être *sans mémoire*. En effet, les épreuves étant indépendantes, la loi de probabilité du nombre d'épreuves à répéter jusqu'à l'obtention d'un premier succès est la même quel que soit le nombre d'échecs obtenus auparavant.

Mathématiquement, cela se traduit par :

$$\forall (n,m) \in \mathbb{N}^{*2},\ P\left(X = n+m \mid X > m\right) = P(X=n)\,,$$

ou, puisque $P(X>n) = \displaystyle\sum_{k=n+1}^{+\infty} P(X=k)$ par :

$$\forall (n,m) \in \mathbb{N}^{*2},\ P\left(X > n+m \mid X > m\right) = P(X>n)\,.$$

La loi géométrique est la seule loi de probabilité discrète qui possède cette propriété.

> ✓ Pour une loi géométrique, l'évènement $(X>n)$ lorsque $n \in \mathbb{N}^*$ représente le fait de n'avoir que des échecs lors des n premières épreuves ; on a donc $P(X>n) = (1-p)^n$ (et cela reste vrai pour $n=0$).
>
> Il est facile de vérifier que l'on a bien : $(1-p)^n = \displaystyle\sum_{k=n+1}^{+\infty} P(X=k)$.

- **Loi de Poisson**
  Soit λ un réel strictement positif. On dit qu'une variable aléatoire sur un espace probabilisé $(\Omega,\mathcal{A},P)$ suit la loi de Poisson de paramètre λ, et l'on note $X \hookrightarrow \mathcal{P}(\lambda)$ si :
  $$\text{(i) } X(\Omega) = \mathbb{N} \qquad \text{(ii) } \forall\, n \in \mathbb{N},\ P(X=n) = \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!}\cdot$$
  Situation : la loi de Poisson sert à estimer le nombre de réalisations d'un évènement dans un intervalle de temps donné (arrivée de clients au guichet d'une banque en une heure, arrivée de malades aux urgences d'un hôpital en une nuit…). Le paramètre λ correspond au nombre moyen d'évènements survenus.
  Ce n'est bien sûr qu'une approximation, qui est justifiée par le résultat suivant.

- **Approximation d'une loi binomiale par une loi de Poisson**
  Si (X_n) est une suite de variables aléatoires telle que X_n suit la loi binomiale $\mathcal{B}(n,p_n)$ avec $\lim\limits_{n\to+\infty} np_n = \lambda > 0$, alors la suite (X_n) converge simplement vers une variable aléatoire X qui suit la loi de Poisson de paramètre λ :
  $$\forall\, k \in \mathbb{N},\ \lim_{n\to+\infty} P(X_n=k) = \lim_{n\to+\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} = P(X=k)\,.$$

  ✓ On considère en général que l'on peut approcher la loi binomiale de paramètre (n,p) par la loi de Poisson de paramètre $\lambda = np$ lorsque $n \geqslant 30$ et $p \leqslant 0{,}1$ avec $np \leqslant 10$. L'avantage de cette approximation est que la loi de Poisson est bien plus rapide à calculer que la loi binomiale.

[S16.4] Couples de variables aléatoires discrètes

- **Loi d'un couple de variables aléatoires**
  Soit $(\Omega,\mathcal{A},P)$, un espace probabilisé et E un ensemble. Soient X et Y deux variables aléatoires discrètes sur $(\Omega,\mathcal{A})$, à valeurs dans E.
  - On appelle loi conjointe du couple de variables aléatoires discrètes (X,Y) la donnée de l'ensemble des couples $\big((x_i,y_j),p_{i,j}\big)$ où
    $$X(\Omega) = \{x_i,\ i \in I\}\,,\ Y(\Omega) = \{y_j,\ j \in J\} \quad (I,\ J \text{ parties de } \mathbb{N})$$
    $$\text{et } p_{i,j} = P\big((X,Y)=(x_i,y_j)\big) = P\left(X = x_i \cap Y = y_j\right)\,.$$
  - Les lois de X et Y s'appellent alors les lois marginales du couple (X,Y).
  - Pour tout indice $j \in J$ tel que $P(Y = y_j)$ soit non nul, on appelle loi conditionnelle de X sachant $(Y=y_j)$ l'application de $X(\Omega)$ dans $[0\,;1]$ définie par la relation :
    $$x_i \longmapsto P_{(Y=y_j)}(X=x_i) = P\left(X=x_i \mid Y=y_j\right) = \frac{P\left(X=x_i \cap Y=y_j\right)}{P(Y=y_j)}\,.$$
  - Pour tout indice $i \in I$ tel que $P(X = x_i)$ soit non nul, on appelle loi conditionnelle de Y sachant $(X=x_i)$ l'application de $Y(\Omega)$ dans $[0\,;1]$ définie par la relation :
    $$y_j \longmapsto P_{(X=x_i)}(Y=y_j) = P\left(Y=y_j \mid X=x_i\right) = \frac{P\left(X=x_i \cap Y=y_j\right)}{P(X=x_i)}\,.$$

- **Lien entre lois conjointes et lois marginales**

  Il est important de bien savoir jongler entre loi du couple, loi marginale et loi conditionnelle. Il s'agit en fait uniquement d'appliquer la définition des probabilités conditionnelles ou alors d'appliquer la formule des probabilités totales. On a donc les relations suivantes.

  . Lois marginales à partir de la loi du couple :

$$P(X = x_i) = \sum_{j \in J} P\left(X = x_i \cap Y = y_j\right)$$
$$P(Y = y_j) = \sum_{i \in I} P\left(X = x_i \cap Y = y_j\right)$$

  . Lois marginales à partir des lois conditionnelles :

$$P(X = x_i) = \sum_{j \in J} P(Y = y_j) P_{(Y=y_j)}(X = x_i)$$
$$P(Y = y_j) = \sum_{i \in I} P(X = x_i) P_{(X=x_i)}(Y = y_j)$$

  . Loi du couple à partir des lois conditionnelles :

$$P\left(X = x_i \cap Y = y_j\right) = P(X = x_i) P_{(X=x_i)}(Y = y_j)$$
$$P\left(X = x_i \cap Y = y_j\right) = P(Y = y_j) P_{(Y=y_j)}(X = x_i)$$

- **Indépendance de variables aléatoires**

  . Soit (X,Y) un couple de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$.

  On dit que les variables X et Y sont indépendantes si :

$$\forall\, i \in I,\, \forall\, j \in J,\, P\left(X = x_i \cap Y = y_j\right) = P(X = x_i) P(Y = y_j)\,.$$

  Cela équivaut à dire que la loi conjointe est le produit des lois marginales.

  . X et Y sont indépendantes si et seulement si :

$$\forall\, A \subset X(\Omega),\, \forall\, B \subset Y(\Omega),\, P\left((X \in A) \cap (Y \in B)\right) = P(X \in A) \times P(Y \in B)\,.$$

  . Soient $X_1,\ldots,X_n$ n variables aléatoires discrètes. On dit que les variables $X_1,\ldots,X_n$ sont mutuellement indépendantes (ou tout simplement indépendantes) lorsque pour tout $(x_1,\ldots,x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$:

$$P\big((X_1 = x_1) \cap \cdots \cap (X_n = x_n)\big) = P(X_1 = x_1) \times \cdots \times P(X_n = x_n)\,.$$

- **Fonction de deux variables aléatoires indépendantes**

  . Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, indépendantes. Soient f et g deux applications définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement.

  Alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

  Le résultat du théorème précédent s'étend à plus de deux variables aléatoires.

- Soient $X_1,\dots,X_n$ n variables aléatoires discrètes mutuellement indépendantes et soit $p \in [\![2\,;n-1]\!]$. Alors toute variable aléatoire fonction des variables $X_1,\dots,X_p$ est indépendante de toute variable aléatoire fonction des variables $X_{p+1},\dots,X_n$.

- **Loi d'une fonction de deux variables aléatoires**

  Soit X et Y deux variables aléatoires discrètes et g une fonction quelconque définie sur $X(\Omega)\times Y(\Omega)$. Alors $Z=g(X,Y)$ est une variable aléatoire discrète.

- **Somme de variables aléatoires**

  - Soient $(\Omega,\mathcal{A},P)$, un espace probabilisé et (X,Y) un couple de variables aléatoires discrètes sur Ω, à valeurs dans $\mathbb{R}$. Notons $X(\Omega) = \{x_i,\, i\in I\}$ et $Y(\Omega)=\{y_j,\, j\in J\}$ où I, J sont deux parties de $\mathbb{N}$.

    Soit $Z=X+Y$. On sait déjà que Z est une variable aléatoire réelle discrète. Pour tout $z \in Z(\Omega)$, l'évènement $(Z=z)$ est la réunion disjointe et au plus dénombrable des évènements $(X=x_i \cap Y=y_j)$ pour tous les couples $(i,j)\in I\times J$ tels que $x_i+y_j=z$. On aura donc :

    $$P(Z=z)=\sum_{(i,j)|x_i+y_j=z} P\left(X=x_i\cap Y=y_j\right).$$

    Dans le cas où les variables X et Y sont indépendantes on aura donc aussi :

    $$P(Z=z)=\sum_{(i,j)|x_i+y_j=z} P(X=x_i)(Y=y_j).$$

  - Si $X_1,\dots,X_n$ sont n variables aléatoires mutuellement indépendantes suivant la loi de Bernoulli de paramètre p, la variable aléatoire $X=X_1+X_2+\cdots+X_n$ suit la loi binomiale $\mathcal{B}(n,p)$.
  - Si $X_1,\dots,X_k$ sont k variables aléatoires mutuellement indépendantes suivant les lois binomiales de paramètres respectifs $(n_1,p),\dots,(n_k,p)$ alors la variable aléatoire $X_1+\cdots+X_k$ suit la loi binomiale de paramètre $\left(\sum_{i=1}^{k} n_i\,,p\right)$.
  - Si X et Y sont deux variables aléatoires indépendantes suivant les lois de Poisson de paramètres respectifs λ et μ alors $X+Y$ suit la loi de Poisson de paramètre $\lambda+\mu$.

[S16.5] Espérance d'une variable aléatoire discrète

Les variables aléatoires considérées dans ce paragraphe sont à valeurs réelles.

- **Définition**

  Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$.

  – Si $X(\Omega)$ est fini et si $(x_i,p_i)_{1\leqslant i\leqslant n}$ est la loi de X, alors l'espérance de X est :

  $$E(X)=\sum_{1\leqslant i\leqslant n} p_i x_i=\sum_{i=1}^{n} x_i P(X=x_i).$$

– Si $X(\Omega)$ est infini ($X(\Omega) = \{x_i,\, i \in I\}$ avec I dénombrable), et si $(x_i,p_i)_{i\in I}$ est la loi de X, alors on dira que X admet une espérance si la famille $(x_ip_i)_{i\in I}$ est sommable. Dans ce cas l'espérance de X est :

$$E(X) = \sum_{i\in I} x_ip_i = \sum_{i\in I} x_iP(X = x_i)\,.$$

Remarque : une variable aléatoire discrète finie a toujours une espérance ce qui n'est pas forcément le cas pour une variable aléatoire discrète infinie.

- **Cas d'une variable aléatoire à valeurs entières**

  Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$.

  X admet une espérance si et seulement si la série de terme général $P(X \geqslant k)$ converge, et dans ce cas :

$$E(X) = \sum_{k=1}^{+\infty} P(X \geqslant k) = \sum_{k=0}^{+\infty} P(X > k)\,.$$

- **Positivité de l'espérance**

  Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$. Si $X \geqslant 0$ (c'est-à-dire si $X(\omega) \geqslant 0$ pour tout $\omega \in \Omega$) et si son espérance existe, alors $E(X) \geqslant 0$.

  De plus, si $X \geqslant 0$ et si $E(X) = 0$, alors $P(X = 0) = 1$ (on dit que X est quasi certaine).

- **Théorème de transfert**

  Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans un ensemble E. Notons $X(\Omega) = \{x_i,\, i \in I\}$ avec I dénombrable.

  Soit f une application de $X(\Omega)$ dans $\mathbb{R}$. La variable aléatoire réelle $f(X)$ admet une espérance si et seulement si la famille $\big(f(x_i)P(X = x_i)\big)_{i\in I}$ est sommable, et dans ce cas :

$$E\big(f(X)\big) = \sum_{i\in I} f(x_i)P(X = x_i)\,.$$

  L'intérêt de ce théorème est de pouvoir calculer l'espérance de $f(X)$ sans avoir besoin de calculer sa loi, seulement en connaissant la loi de X.

- **Linéarité de l'espérance**

  Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, et soit $\lambda \in \mathbb{R}$.

  Si les espérances de X et de Y existent, alors l'espérance de $\lambda X + Y$ existe et

$$E(\lambda X + Y) = \lambda E(X) + E(Y)\,.$$

  Soit X une variable aléatoire discrète dont l'espérance existe. Puisque $E(1) = 1$ il résulte de l'égalité précédente que la variable aléatoire $X - E(X)$ est d'espérance nulle. On l'appelle variable aléatoire centrée associée à X.

- **Croissance de l'espérance**
  - Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, dont l'espérance existe.
    Si $X \leqslant Y$ sur Ω, alors $E(X) \leqslant E(Y)$.
  - Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$.
    On suppose que l'espérance de Y existe, et que $|X| \leqslant Y$ sur Ω.
    Alors l'espérance de X existe.
    Cas particulier : si X est une variable aléatoire discrète bornée, son espérance existe.

- **Produit de deux variables aléatoires indépendantes**
  Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, dont l'espérance existe.
  Si X et Y sont indépendantes, la variable aléatoire $Z = XY$ possède une espérance finie et l'on a :
  $$E(XY) = E(X)E(Y)\,.$$

- **Inégalité de Markov**
  Soit X une variable aléatoire discrète réelle positive, dont l'espérance existe.
  Alors, pour tout réel $a > 0$:
  $$P(X \geqslant a) \leqslant \frac{E(X)}{a}\,.$$

[S16.6] Moments d'une variable aléatoire réelle discrète

Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé, X une variable aléatoire réelle discrète et soit $r \geqslant 1$ un entier.

– Si $X(\Omega)$ est fini et si $(x_i,p_i)_{1\leqslant i\leqslant n}$ est la loi de X, alors le moment d'ordre r de X est :
$$m_r(X) = E(X^r) = \sum_{i=1}^{n} p_i x_i^r\,.$$

– Si $X(\Omega)$ est infini et si $(x_i,p_i)_{i\in I}$ est la loi de X, alors on dira que X admet un moment d'ordre r si la famille $(x_i^r p_i)_{i\in I}$ est sommable. Dans ce cas le moment d'ordre r de X est :
$$m_r(X) = E(X^r) = \sum_{i\in I} x_i^r p_i = \sum_{i\in I} x_i^r P(X = x_i)\,.$$

Théorème
Si X admet un moment d'ordre $r \geqslant 2$, elle admet des moments d'ordre s pour tout $s \in [\![1\,;r]\!]$.

[S16.7] Variance

- Soit X une variable aléatoire réelle dont le moment d'ordre 2, c'est-à-dire $E(X^2)$, existe.
  - La variable aléatoire $X - E(X)$ admet aussi un moment d'ordre 2, appelé variance de X. Ainsi :
  $$V(X) = E\left(\left(X - E(X)\right)^2\right).$$
  - On a la relation :
  $$V(X) = E(X^2) - E(X)^2 \quad (\textit{formule de Kœnig-Huygens}).$$

  On appelle alors écart -type de X le réel $\sigma(X) = \sqrt{V(X)}$.
- Soit X une variable aléatoire réelle dont la variance $V(X)$ existe.
  Si a et b sont deux réels, la variable $aX + b$ admet aussi une variance et :
  $$V(aX + b) = a^2 V(X).$$
- **Inégalité de Bienaymé-Tchebychev**
  Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Alors, pour tout réel $\varepsilon > 0$:
  $$P\big(|X - E(X)| \geqslant \varepsilon\big) \leqslant \frac{V(X)}{\varepsilon^2}.$$
  Cette inégalité montre que la variance permet de mesurer la dispersion de la variable X autour de sa moyenne.
  Conséquence
  Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Notons $m = E(X)$.
  Si $V(X) = 0$ alors $P(X = m) = 1$ (X est quasi certaine).

[S16.8] Moments des lois usuelles

- **Loi uniforme**
  Soit $X \hookrightarrow \mathcal{U}(\llbracket 1\,;n \rrbracket$. Alors : $E(X) = \dfrac{n+1}{2}$ et $V(X) = \dfrac{n^2-1}{12}$.
- **Loi binomiale**
  Soit $X \hookrightarrow \mathcal{B}(n,p)$. Alors : $E(X) = np$ et $V(X) = npq$.
- **Loi géométrique**
  Soit $X \hookrightarrow \mathcal{G}(p)$. Alors : $E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{q}{p^2}$.
- **Loi de Poisson**
  Soit $X \hookrightarrow \mathcal{P}(\lambda)$. Alors : $E(X) = \lambda$ et $V(X) = \lambda$.

[S16.9] Covariance

- Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega,\mathcal{A},P)$, ayant chacune un moment d'ordre 2.
  Alors l'espérance de XY existe et

$$E(XY)^2 \leqslant E(X^2)E(Y^2) \quad (\textit{inégalité de Cauchy-Schwarz}).$$

- Si X et Y sont deux variables aléatoires réelles admettant un moment d'ordre 2, alors la variable aléatoire $X+Y$ admet aussi un moment d'ordre 2.

- Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2, on appelle covariance de X et Y le réel :

$$\text{cov}\,(X,Y) = E\big[(X-E(X))(Y-E(Y))\big]$$

  (l'existence est assurée par les propriétés précédentes).
  Un calcul simple montre que l'on a aussi : $\text{cov}\,(X,Y) = E(XY) - E(X)E(Y)$.
  On a alors la relation :

$$V(X+Y) = V(X) + V(Y) + 2\text{cov}\,(X,Y).$$

- **Propriétés**
  Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé.
  - L'ensemble $\mathcal{V}(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω est un sous-espace vectoriel de $\mathcal{A}(\Omega,\mathbb{R})$.
  - L'ensemble $\mathcal{V}^1(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant une espérance est un sous-espace vectoriel de $\mathcal{V}(\Omega,\mathbb{R})$.
  - L'ensemble $\mathcal{V}^2(\Omega,\mathbb{R})$ des variables aléatoires réelles discrètes sur Ω admettant un moment d'ordre 2 est un sous-espace vectoriel de $\mathcal{V}^1(\Omega,\mathbb{R})$.
  - L'application $(X,Y) \mapsto \text{cov}\,(X,Y)$ est une forme bilinéaire symétrique positive (mais non définie) sur l'espace vectoriel $\mathcal{V}^2(\Omega,\mathbb{R})$, et pour tout $X \in \mathcal{V}^2(\Omega,\mathbb{R})$, $\text{cov}\,(X,X) = V(X)$.
  - Si X et Y appartiennent à $\mathcal{V}^2(\Omega,\mathbb{R})$, on a l'inégalité de Cauchy-Schwarz :

$$|\text{cov}\,(X,Y)| \leqslant \sigma(X)\sigma(Y).$$

- Si X et Y sont deux variables aléatoires réelles admettant chacune un moment d'ordre 2 et **indépendantes**, alors leur covariance est nulle.
  Il en résulte que l'on a alors $V(X+Y) = V(X) + V(Y)$.

  ✓ La réciproque de cette propriété est fausse : deux variables aléatoires peuvent avoir une covariance nulle sans être indépendantes.

- **Généralisation**
  Si $X_1,\ldots,X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2, en utilisant la bilinéarité de la covariance on obtient :

$$V(X_1+\cdots+X_n) = V(X_1)+\cdots+V(X_n) + 2\sum_{1\leqslant i<j\leqslant n} \text{cov}\,(X_i,X_j).$$

Il en résulte que si $X_1, \dots, X_n$ sont n variables aléatoires réelles possédant toutes un moment d'ordre 2 et *indépendantes deux à deux* :

$$V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i)\,.$$

- Soient X et Y deux variables aléatoires réelles dont la variance existe et est non nulle. On appelle coefficient de corrélation linéaire de X et Y le nombre réel :

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)}\,.$$

Propriétés

Soient X et Y deux variables aléatoires réelles dont la variance existe et est non nulle.

- $|\rho(X,Y)| \leqslant 1$;
- $|\rho(X,Y)| = 1$ si et seulement si Y est une fonction quasi affine de X, c'est-à-dire s'il existe a et b réels tels que l'évènement $(Y = aX + b)$ soit quasi certain.

[S16.10] Variables aléatoires à valeurs dans $\mathbb{N}$

- Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$.

  La fonction génératrice de X est la série entière :

$$G_X(t) = \sum_{n=0}^{+\infty} P(X = n)t^n\,.$$

- La fonction génératrice d'une variable aléatoire à valeurs entières a un rayon de convergence R_X supérieur ou égal à 1.
- G_X est de classe $\mathscr{C}^\infty$ sur $]-R_X\,;R_X[$ et :

$$\forall n \in \mathbb{N},\ P(X = n) = \frac{G_X^{(n)}(0)}{n!}\,.$$

- Pour tout $t \in]-R_X\,;R_X[$, $G_X(t) = E(t^X)$.
- **Fonctions génératrices des lois usuelles**
  - Si X suit la loi binomiale de paramètre (n,p), sa fonction génératrice est la fonction polynomiale définie par :

$$\forall t \in \mathbb{R},\ G_X(t) = (pt + q)^n.$$

  - Si X suit la loi géométrique de paramètre p, le rayon de convergence de sa série génératrice est $\dfrac{1}{q}$ et :

$$\forall t \in \left]-\frac{1}{q}\,;\frac{1}{q}\right[,\ G_X(t) = \frac{pt}{1-qt}\,.$$

- Si X suit la loi de Poisson de paramètre λ, sa série génératrice est de rayon de convergence $+\infty$ et :
$$\forall t \in \mathbb{R},\ G_X(t) = \mathrm{e}^{\lambda(t-1)}.$$

- **Somme de variables aléatoires indépendantes**
  - Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$, et G_X, G_Y leurs fonctions génératrices respectives.
    Si X et Y sont indépendantes on a :
$$\forall t \in [-1\,;1],\ \ G_{X+Y}(t) = G_X(t)G_Y(t)\,.$$
  - Plus généralement, si $X_1,\ldots,X_n$ sont n variables aléatoires mutuellement indépendantes, on a :
$$\forall t \in [-1\,;1],\ G_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} G_{X_i}(t)\,.$$

- **Fonction génératrice et moments**
  - Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice. On suppose que $R_X > 1$, ce qui implique que G_X est de classe $\mathscr{C}^\infty$ au voisinage de 1. Alors :
$$\forall k \in \mathbb{N}^*,\ G_X^{(k)}(1) = E\big(X(X-1)\ldots(X-k+1)\big)\,.$$
    En particulier :
$$E(X) = G'_X(1)\,,\ E(X(X-1)) = G''_X(1) \text{ d'où } E(X^2) = G'_X(1) + G''_X(1)\,.$$
  - Réciproquement, soit X une variable aléatoire définie sur un espace probabilisé $(\Omega,\mathcal{A},P)$, à valeurs dans $\mathbb{N}$, et G_X sa fonction génératrice.
    L'espérance de X existe si et seulement si G_X est dérivable à gauche en 1 et la variance de X existe si et seulement si G_X est deux fois dérivable à gauche en 1.
    Dans ce cas on aura : $E(X) = G'_X(1^-)$ et $E(X^2) = G'_X(1^-) + G''_X(1^-)$.

[S16.11] Loi faible des grands nombres

Soit $(X_i)_{i\in N}$ une suite de variables aléatoires réelles définies sur un espace probabilisé $(\Omega,\mathcal{A},P)$ indépendantes, de même loi, possédant une espérance $m = E(X)$ et une variance $\sigma^2 = V(X)$. On pose :
$$\overline{X_n} = \frac{\displaystyle\sum_{i=1}^{n} X_i}{n}\,.$$
Alors on a :
$$\forall \varepsilon > 0,\ P\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}\,.$$
Par conséquent :
$$\forall \varepsilon > 0,\ \lim_{n\to+\infty} P\left(\left|\overline{X_n} - m\right| \geqslant \varepsilon\right) = 0\,.$$

Thème 17 - Équations différentielles linéaires

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, I est un intervalle de $\mathbb{R}$ et E est un $\mathbb{K}$-espace vectoriel normé de dimension finie $n \in \mathbb{N}^*$.

[S17.1] Équations différentielles linéaires, systèmes différentiels

- Une équation différentielle (vectorielle) linéaire du premier ordre est une équation de la forme :
$$x' = a(x) + b \qquad (L)$$
où a est une application continue de I dans $\mathscr{L}(E)$ et b une application continue de I dans E.
Une solution de cette équation est une application x dérivable de I dans E telle que :
$$\forall t \in I,\ x'(t) = a(t)\big(x(t)\big) + b(t)\,.$$

  ✓ Dans cette écriture, $a(t)$ est un endomorphisme de E et $a(t)\big(x(t)\big)$ représente donc l'image du vecteur $x(t)$ par cet endomorphisme.

- Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E, et :
  - $X(t) = {}^{\mathrm{t}}\begin{pmatrix} x_1(t) & \ldots & x_n(t)\end{pmatrix}$ la matrice colonne formée des coordonnées du vecteur $x(t)$ dans $\mathscr{B}$, où les x_i sont des applications dérivables de I dans $\mathbb{K}$;
  - $B(t) = {}^{\mathrm{t}}\begin{pmatrix} b_1(t) & \ldots & b_n(t)\end{pmatrix}$ la matrice colonne formée des coordonnées du vecteur $b(t)$ dans $\mathscr{B}$, où les b_i sont des applications continues de I dans $\mathbb{K}$;
  - $A(t) = \big(a_{ij}(t)\big)_{1 \leqslant i,j \leqslant n}$ la matrice dans $\mathscr{B}$ de l'application linéaire $a(t)$, où les a_{ij} sont des applications continues de I dans $\mathbb{K}$.

  L'équation $x' = a(x) + b$ peut alors s'écrire matriciellement sous la forme :
$$X' = AX + B \quad \text{ou :} \quad \forall t \in I\ ,\ X'(t) = A(t)X(t) + B(t) \qquad (L)$$
  On obtient alors ce que l'on appelle un système différentiel :
$$\begin{cases} x_1'(t) &= a_{11}(t)x_1(t) + \cdots + a_{1n}(t)x_n(t) + b_1(t) \\ x_2'(t) &= a_{21}(t)x_1(t) + \cdots + a_{2n}(t)x_n(t) + b_2(t) \\ \vdots & \qquad\qquad \vdots \\ x_n'(t) &= a_{n1}(t)x_1(t) + \cdots + a_{nn}(t)x_n(t) + b_n(t) \end{cases}$$

[S17.2] Théorème de Cauchy-Lipschitz linéaire

- Résoudre le problème de Cauchy pour une équation différentielle $x' = a(x)+b$ sur I consiste à étudier l'existence et l'unicité d'une solution vérifiant une condition initiale du type $x(t_0) = x_0$ avec $t_0 \in I$ et $x_0 \in E$ donnés.

- **Théorème de Cauchy-Lipschitz linéaire**
  Soit E un espace vectoriel de dimension finie, I un intervalle de $\mathbb{R}$, a une application continue de I dans $\mathscr{L}(E)$ et b une application continue de I dans E.
  Pour tout $(t_0,x_0) \in I \times E$, il existe une et une seule solution sur I de l'équation différentielle $x' = a(x) + b$ vérifiant la condition initiale $x(t_0) = x_0$.

  ✓ La continuité des applications a et b est indispensable.

  Par exemple, l'équation différentielle $x' = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$ n'a pas de solution sur $\mathbb{R}$ (en effet, une telle solution ne pourrait être dérivable en 0).

- Matriciellement, ce théorème s'énonce ainsi.
  Soient $A\colon I \to \mathcal{M}_n(\mathbb{K})$ et $B\colon I \to \mathcal{M}_{n,1}(\mathbb{K})$ des applications continues.
  Soit $t_0 \in I$, et soit $X_0 \in \mathcal{M}_{n,1}(\mathbb{K})$.
  Alors il existe une unique application $X\colon I \to \mathcal{M}_{n,1}(\mathbb{K})$ telle que :
  $$\forall t \in I,\ X'(t) = A(t)X(t) + B(t) \quad \text{et} \quad X(t_0) = X_0\,.$$

[S17.3] Structure de l'ensemble des solutions

- **Solutions d'une équation homogène**
  L'équation (H) : $x' = a(x)$ s'appelle l'équation homogène associée à l'équation (L) : $x' = a(x) + b$.
  L'ensemble $\mathscr{S}_H$ des solutions de l'équation différentielle homogène (H) est un sous-espace vectoriel de l'espace vectoriel $\mathscr{C}^1(I,E)$ des fonctions de classe $\mathscr{C}^1$ de I dans E.

  De plus, cet espace vectoriel est de dimension finie, et $\dim \mathscr{S}_H = \dim E$.

  ✓ Ce résultat n'est valable que sur un *intervalle*.

  Exemple : l'espace des solutions sur $\mathbb{R}^*$ de l'équation $x' = 0$ est de dimension 2, puisque les solutions sont les applications telles que $x(t) = C_1$ sur $]-\infty\,;0[$ et $x(t) = C_2$ sur $]0\,;+\infty[$.

- **Solutions d'une équation avec second membre**
  Si x_L est une solution particulière de l'équation (L) : $x' = a(x) + b$ (il en existe, d'après le théorème de Cauchy-Lipschitz), alors l'ensemble des solutions de l'équation (L) est exactement l'ensemble des fonctions somme de x_L et d'une solution quelconque de l'équation homogène (H) :
  $$\mathscr{S}_L = x_L + \mathscr{S}_H\,.$$

- **Principe de superposition des solutions**
  Soient $b_1 : I \to E$ et $b_2 : I \to E$ deux applications continues, et λ_1, λ_2 dans $\mathbb{K}$.
  Soient x_1 et x_2 deux solutions respectives des équations différentielles
  $$(L_1) : \quad x' = a(x) + b_1 \qquad \text{et} \qquad (L_2) : \quad x' = a(x) + b_2\,.$$
  Alors $\lambda_1 x_1 + \lambda_2 x_2$ est solution de l'équation différentielle :
  $$(L) : \quad x' = a(x) + \lambda_1 b_1 + \lambda_2 b_2\,.$$

- **Système fondamental de solution, wronskien**

  - Soit $\mathscr{B}$ une base de E, et $(x_1,\dots,x_n)$ une famille formée de n solutions de l'équation homogène (H).

    On appelle matrice wronskienne de cette famille la matrice $\mathcal{W}(t)$ des vecteurs $(x_1(t),\dots,x_n(t))$ dans la base $\mathscr{B}$ et on appelle wronskien de cette famille, noté $W(t)$, le déterminant de cette matrice :

    $$\forall\, t\in I,\ W(t)=\det\mathcal{W}(t)=\det_{\mathscr{B}}\big(x_1(t),\dots,x_n(t)\big).$$

  - On appelle système fondamental de solutions de l'équation $(L)\ : x'=a(x)+b$ toute base de l'espace vectoriel $\mathscr{S}_H$ des solutions de l'équation homogène (H).

    Les propositions suivantes sont équivalentes :

    (i) $(x_1,\dots,x_n)$ est un système fondamental de solutions de (H),

    (ii) $\forall\, t\in I,\ W(t)\neq 0$,

    (iii) $\exists\, t_0\in I$ tel que $W(t_0)\neq 0$.

  - Si $(x_1,\dots,x_n)$ est un système fondamental de solutions, toute solution de l'équation homogène $x'=a(x)$ est donc combinaison linéaire de $(x_1,\dots,x_n)$:

    $$\mathscr{S}_H=\mathrm{Vect}\big(x_1,\dots,x_n\big)=\left\{\sum_{k=1}^{n}\lambda_k x_k\ \middle|\ (\lambda_1,\dots,\lambda_n)\in\mathbb{K}^n\right\}.$$

    En notant $X(t)$ la matrice colonne des coordonnées de $x(t)$ dans $\mathscr{B}$, la relation $x=\sum_{k=1}^{n}\lambda_k x_k$ s'écrit $X(t)=\mathcal{W}(t)C$ où $\mathcal{W}$ est la matrice wronskienne de $(x_1,\dots,x_n)$ et C le vecteur colonne $C={}^{\mathrm{t}}\begin{pmatrix}\lambda_1 & \dots & \lambda_n\end{pmatrix}$, de sorte que, matriciellement :

    $$\mathscr{S}_H=\{t\mapsto\mathcal{W}(t)C\mid C\in\mathcal{M}_{n,1}(\mathbb{K})\}\,.$$

[S17.4] Méthode de Lagrange, ou de variation des constantes

On considère ici une équation différentielle linéaire écrite matriciellement dans une base $\mathscr{B}$ de E sous la forme :

$$(L)\quad X'(t)=A(t)X(t)+B(t)\,.$$

On suppose qu'on connaît un système fondamental de solutions $(x_1,\dots,x_n)$, et on note $\mathcal{W}$ sa matrice wronskienne dans $\mathscr{B}$.

La méthode de Lagrange consiste à chercher les solutions de (L) sous la forme $X(t)=\mathcal{W}(t)C(t)$ où C est une application de classe $\mathscr{C}^1$ de I dans $\mathcal{M}_{n,1}(\mathbb{K})$.

Plus précisément :

$$\begin{aligned}X(t)=\mathcal{W}(t)C(t)\text{ est solution de }(L)&\iff\mathcal{W}(t)C'(t)=B(t)\\&\iff C'(t)=\mathcal{W}(t)^{-1}B(t)\,.\end{aligned}$$

Si $C(t) = {}^{\mathrm{t}}\big(\lambda_1(t) \quad \dots \quad \lambda_n(t)\big)$, cela revient à chercher les solutions de (L) sous la forme :

$$x(t) = \sum_{k=1}^{n} \lambda_k(t) x_k(t)\,,$$

avec la condition : $\displaystyle\sum_{k=1}^{n} \lambda_k'(t) x_k(t) = b(t)$. La résolution de ce système permet alors d'obtenir les fonctions λ_k' puis les fonctions λ_k (à une constante près).

[S17.5] Systèmes différentiels linéaires à coefficients constants

On s'intéresse ici à des systèmes différentiels de la forme : $X'(t) = AX(t) + B(t)$, où $A \in \mathcal{M}_n(\mathbb{K})$ ne dépend pas de t.

- **Utilisation de l'exponentielle de matrice**

  - **Rappels**

    – Nous avons vu en **[S11.9]** que l'on peut définir l'exponentielle d'une matrice carrée A, ou d'un endomorphisme a d'un espace vectoriel de dimension finie, par :

    $$\mathrm{e}^A = \exp(A) = \sum_{k=0}^{+\infty} \frac{A^k}{k!} \quad \text{ou} \quad \mathrm{e}^a = \exp(a) = \sum_{k=0}^{+\infty} \frac{a^k}{k!}\,.$$

    De plus, si A, B sont deux matrices carrées qui commutent, ou si a, b sont deux endomorphismes qui commutent, on a :

    $$\exp(A+B) = \exp(A) + \exp(B) \quad \text{ou} \quad \exp(a+b) = \exp(a) \circ \exp(b)\,.$$

    – Avec les mêmes notations, en appliquant les résultats de **[S12.11]** concernant les séries de fonctions de classe $\mathscr{C}^1$, on montre que les applications $\mathrm{e}_A : t \in \mathbb{R} \mapsto \exp(tA)$ ou $\mathrm{e}_a : t \in \mathbb{R} \mapsto \exp(ta)$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et que :

    $$\mathrm{e}_A' = A\mathrm{e}_A = \mathrm{e}_A A \quad \text{ou} \quad \mathrm{e}_a' = a \circ \mathrm{e}_a = \mathrm{e}_a \circ a\,.$$

  - L'ensemble des solutions du système homogène (H) : $X'(t) = AX(t)$ où $A \in \mathcal{M}_n(\mathbb{K})$ est constante est l'ensemble :

    $$\mathscr{S}_H = \left\{ t \mapsto \mathrm{e}^{tA} C \;\middle|\; C \in \mathcal{M}_{n,1}(\mathbb{K}) \right\}.$$

    Autrement dit, un système fondamental de solutions de (H) est formé des vecteurs colonnes de la matrice e^{tA}.

  - Pour résoudre alors l'équation avec second membre (L) : $X'(t) = AX(t) + B(t)$, on utilise la méthode de variations des constantes en cherchant les solutions sous la forme :

    $$X(t) = \mathrm{e}^{tA} C(t)$$

    qui conduit ici à :

    $$C'(t) = \mathrm{e}^{-tA} B(t)\,.$$

- **Utilisation de la réduction**

  . On suppose que A admet une valeur propre λ.

  Si V est un vecteur propre associé à cette valeur propre, alors la fonction $t \mapsto \mathrm{e}^{\lambda t}V$ est solution sur $\mathbb{R}$ de l'équation homogène $X'(t) = AX(t)$.

  . On suppose ici que A est diagonalisable. Il existe donc une base $(V_1, \ldots, V_n)$ de $\mathcal{M}_{n,1}(\mathbb{K})$ formée de vecteurs propres de A, pour les valeurs propres $(\lambda_1, \ldots, \lambda_n)$.

  Alors les fonctions $X_i : t \mapsto \mathrm{e}^{\lambda_i t}V_i$, pour $i \in [\![1\,;n]\!]$, forment une base de l'espace vectoriel $\mathscr{S}_H$ des solutions de l'équation homogène $X'(t) = AX(t)$.
  Toute solution de l'équation homogène est donc de la forme :

$$t \longmapsto \sum_{i=1}^{n} \alpha_i \mathrm{e}^{\lambda_i t}V_i \quad \text{avec } (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n\,.$$

  . De façon plus générale, si la matrice A est semblable à une matrice T plus simple (triangulaire supérieure en général), on effectuera un changement de base, ce qui revient à changer de fonctions inconnues, pour obtenir un système différentiel plus facile à résoudre.

  Le principe est simple : si $T = P^{-1}AP$, le système $X'(t) = AX(t)$ équivaut à $P^{-1}X'(t) = TP^{-1}X(t)$, donc, en posant $Y(t) = P^{-1}X(t)$, soit $X(t) = PY(t)$, on obtiendra le nouveau système différentiel $Y'(t) = TY(t)$ (en effet, $X(t) = PY(t)$ implique $X'(t) = PY'(t)$ car la matrice P est constante).

  ✓ Remarquez que dans cette méthode, dans le cas d'une équation homogène, le calcul de la matrice inverse P^{-1} n'est pas utile en pratique puisque la solution finale s'obtient par $X(t) = PY(t)$.

  ✓ D'après les calculs exposés ci-dessus, la méthode peut aussi s'appliquer dans le cas d'une équation de la forme $X'(t) = A(t)X(t)$ sous réserve que la matrice de passage P ne dépende pas de t.

[S17.6] Équations différentielles linéaires scalaires du premier ordre

Il s'agit d'équations différentielles de la forme :

$$\alpha(t)x' + \beta(t)x = \gamma(t)\,,$$

où les fonctions α, β et γ sont continues sur un certain intervalle I, à valeurs dans $\mathbb{K}$. Les solutions x sont à chercher parmi les applications de I dans $\mathbb{K}$.

- La *première* chose à faire est d'examiner si la fonction α peut ou non s'annuler sur I. Si oui, on cherche les sous-intervalles de I où α ne s'annule pas, et on résoudra l'équation sur *chacun* de ces intervalles (et il y aura des constantes d'intégration différentes sur chaque intervalle).
- Un tel intervalle J étant choisi, on peut alors diviser l'équation par $\alpha(t)$ et on se ramène ainsi à une équation différentielle sous forme résolue :

$$x' = a(t)x + b(t) \quad (L)$$

  On peut remarquer que les solutions d'une telle équation sont nécessairement de classe $\mathscr{C}^1$ sur J.

- **Intégration de l'équation homogène** (H) : $x' = a(t)x$

  Les solutions de l'équation différentielle $x' = a(t)x$ sont les fonctions de la forme :

  $$t \longmapsto C\exp\left(\int_{t_0}^{t} a(u)\,\mathrm{d}u\right),$$

  où C est une constante dans $\mathbb{K}$ et où $t \mapsto \displaystyle\int_{t_0}^{t} a(u)\,\mathrm{d}u$ est une primitive de la fonction a.

  ✓ On retrouve ainsi le théorème de Cauchy-Lipschitz : l'ensemble des solutions de (H) sur J est ici un $\mathbb{K}$-espace vectoriel de dimension 1.

- **Intégration de l'équation avec second membre** (L) : $x' = a(t)x + b(t)$

  L'ensemble des solutions de l'équation complète (L) est l'ensemble des fonctions de la forme $x = x_L + y$, où y est solution de l'équation homogène (H) et où x_L est une solution particulière de (L).

  Si l'on connaît une solution particulière x_L de (L), la résolution est terminée.

  Sinon, on peut rechercher une telle solution par la *méthode de variation de la constante*, expliquée ci-dessous.

  Soit y une solution non nulle de (H), de la forme $y(t) = \exp\left(\displaystyle\int_{t_0}^{t} a(u)\,\mathrm{d}u\right)$. Une telle solution ne peut s'annuler, et on peut donc chercher les solutions de (L) sous la forme $x(t) = C(t)y(t)$ avec C de classe $\mathscr{C}^1$. Un calcul simple montre alors que (L) est équivalente à l'équation :

  $$C'(t)y(t) = b(t) \quad \text{soit} \quad C'(t) = \frac{b(t)}{y(t)}.$$

  Cela permet d'obtenir $C(t)$ par une simple primitivation, puis d'en déduire la solution $x(t)$.

  ✓ Il est bon de retenir les formules ci-dessus, en particulier de se rappeler que dans l'équation transformée apparaît seulement $C'(t)$ et non plus $C(t)$. Il est en fait inutile de refaire le calcul de la dérivée de $t \mapsto C(t)y(t)$ lorsqu'on connaît ces formules.

- **Cas d'une équation à coefficients constants**

  Il s'agit ici d'une équation de la forme $x' - ax = b(t)$, où $a \in \mathbb{K}$ est une constante. Dans ce cas, l'ensemble des solutions de l'équation homogène $x' - ax = 0$ est l'ensemble des fonctions de la forme $t \mapsto Ce^{at}$.

  La méthode de variation de la constante s'applique bien sûr pour la résolution de l'équation complète, mais il y a un cas particulier à retenir.

  Si le second membre est de la forme $b(t) = \mathrm{e}^{mt}P(t)$ où $m \in \mathbb{C}$ et où P est un polynôme, on peut rechercher (avec des coefficients à déterminer) une solution particulière de l'équation $x' - ax = \mathrm{e}^{mt}P(t)$ sous la forme :

  - $\mathrm{e}^{mt}Q(t)$ avec Q polynôme de même degré que P si $m \neq a$,
  - $\mathrm{e}^{mt}tQ(t)$ avec Q polynôme de même degré que P si $m = a$.

- Pour résoudre complètement l'équation initiale $\alpha(t)x' + \beta(t)x = \gamma(t)$ sur tout l'intervalle I, il faut ensuite (éventuellement) raccorder les solutions trouvées précédemment sur les intervalles où la fonction α ne s'annule pas.

[S17.7] Équations différentielles linéaires scalaires du second ordre

Il s'agit d'équations différentielles de la forme :

$$\alpha(t)x'' + \beta(t)x' + \gamma(t)x = \delta(t)\,,$$

où les fonctions α, β, γ et δ sont continues sur un certain intervalle I, à valeurs dans $\mathbb{K}$.

- Comme pour les équations du premier ordre, il faut résoudre une telle équation sur un intervalle $J \subset I$ où la fonction α ne s'annule pas.
  Un tel intervalle J étant choisi, on peut alors diviser l'équation par $\alpha(t)$ et on se ramène ainsi à une équation différentielle sous forme résolue :

  $$x'' = a(t)x' + b(t)x + c(t) \quad (L)$$

  où a, b et c sont des applications continues de J dans $\mathbb{K}$.
  On peut remarquer que les solutions x d'une telle équation sont nécessairement de classe $\mathscr{C}^2$ sur J.

- **Système différentiel associé**
  L'équation (L) précédente peut s'écrire :

  $$\begin{pmatrix} x'(t) \\ x''(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ b(t) & a(t) \end{pmatrix} \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} + \begin{pmatrix} 0 \\ c(t) \end{pmatrix},$$

  soit :

  $$X' = A\,X + B\,,$$

  avec : $X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix}$, $A(t) = \begin{pmatrix} 0 & 1 \\ b(t) & c(t) \end{pmatrix}$ et $B(t) = \begin{pmatrix} 0 \\ c(t) \end{pmatrix}$.

  On est donc ramené à un système différentiel linéaire du premier ordre. Les résultats obtenus pour ce type d'équation permettent alors d'obtenir les résultats qui suivent.

- Une famille (x_1,x_2) de solutions de l'équation homogène : $x'' = a(t)x' + b(t)x$ est un système fondamental de solutions si et seulement si le wronskien $W(t) = \begin{vmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{vmatrix}$ est non nul en tout point de J (il suffit qu'il le soit en au moins un point).

  Dans ce cas, l'ensemble $\mathscr{S}_H$ des solutions de l'équation homogène est l'espace vectoriel engendré par x_1 et x_2 soit :

  $$\mathscr{S}_H = \left\{\lambda_1 x_1 + \lambda_2 x_2 \ ,\ (\lambda_1,\lambda_2) \in \mathbb{K}^2\right\}.$$

- L'ensemble $\mathscr{S}_L$ des solutions de (L) sur J est un espace affine de dimension 2 : $\mathscr{S}_L = x_L + \mathscr{S}_H$, où x_L est une solution particulière de (L).

- Le théorème de Cauchy-Lipschitz linéaire s'énonce ici : pour tout $t_0 \in J$ et tout $(x_0,x_0') \in \mathbb{K}^2$, il existe une solution et une seule solution de (L) sur J vérifiant les conditions initiales :

  $$x(t_0) = x_0 \quad \text{et} \quad x'(t_0) = x_0'\,.$$

- **Cas où l'on connaît une solution de l'équation homogène**

  Dans ce cas, on peut, comme pour les équations du premier ordre, appliquer la méthode de variation de la constante (encore appelée *méthode de Lagrange*). En voici le principe.

  Considérons l'équation différentielle (pas forcément écrite sous forme résolue) :

  $$\alpha(t)x'' + \beta(t)x' + \gamma(t)x = \delta(t) \quad (L)\,.$$

  Supposons connue une solution x_L de l'équation *homogène*.

  Sur un intervalle où la fonction x_L ne s'annule pas, on peut chercher une solution quelconque de l'équation (L) sous la forme : $x(t) = y(t)x_L(t)$ (en effet, puisque x_L ne s'annule pas, la fonction $y = \dfrac{x}{x_L}$ est bien définie et deux fois dérivable).

  Quelques calculs montrent alors que x est solution de (L) si et seulement si :

  $$\alpha(t)x_L(t)y''(t) + (2\alpha(t)x_L'(t) + \beta(t)x_L(t))\,y'(t) = \delta(t)\,.$$

  C'est une équation scalaire du premier ordre pour la fonction y'. Il ne reste « plus qu'à » la résoudre pour trouver y', d'où l'on déduit y puis x.

- **Cas où l'on connaît deux solutions de l'équation homogène**

  Considérons l'équation différentielle :

  $$x'' = a(t)x' + b(t)x + c(t) \quad (L)\,.$$

  On suppose connu un système fondamental de solutions (x_1,x_2) de l'équation homogène (H) : $x'' = a(t)x' + b(t)x$.

  La méthode de variations des constantes consiste ici à chercher les solutions x de (L) sous la forme :

  $$x(t) = \lambda_1(t)x_1(t) + \lambda_2(t)x_2(t)\,,$$

  où λ_1, λ_2 sont de classe $\mathscr{C}^1$ et vérifient :

  $$\lambda_1'(t)\begin{pmatrix}x_1(t)\\x_1'(t)\end{pmatrix} + \lambda_2'(t)\begin{pmatrix}x_2(t)\\x_2'(t)\end{pmatrix} = \begin{pmatrix}0\\c(t)\end{pmatrix}\,,$$

  ce qui amène à résoudre le système :

  $$\begin{cases}\lambda_1'x_1 + \lambda_2'x_2 = 0\\ \lambda_1'x_1' + \lambda_2'x_2' = c\end{cases}\,.$$

  Ce système permet de calculer λ_1' et λ_2', d'où on déduit par primitivation λ_1 et λ_2, et finalement la solution générale de (L) : $x = \lambda_1x_1 + \lambda_2x_2$.

  > ✓ Il ne sert à rien d'apprendre ces formules par cœur. La meilleure façon de retenir cette méthode consiste à bien comprendre d'où elle provient : il s'agit simplement de la méthode de variation des constantes pour le système différentiel associé à l'équation (L).

[S17.8] Équations différentielles du 2e ordre à coefficients constants

Soit (L) une équation différentielle de la forme :

$$x'' + ax' + bx = c(t) \quad \text{avec } a,\, b \in \mathbb{K} \text{ et } c \text{ continue sur } I \text{ à valeurs dans } \mathbb{K}$$

(où $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$).

L'équation homogène associée est $(H) : x'' + ax' + bx = 0$.

- **Résolution de l'équation homogène**

  On cherche des solutions de cette équation de la forme $t \mapsto \mathrm{e}^{rt}$. Un calcul immédiat montre que cette fonction est solution de (H) si et seulement si $r^2 + ar + b = 0$.

  À l'équation (H), on associe donc son *équation caractéristique* :

  $$(E_c) : \; r^2 + ar + b = 0\,,$$

  et on étudie ses solutions. Il faut distinguer deux cas, selon que le corps de base est $\mathbb{R}$ ou $\mathbb{C}$; on notera Δ le discriminant de l'équation caractéristique.

  - Si $\mathbb{K} = \mathbb{C}$
    - Si $\Delta \neq 0$, (E_c) possède deux racines distinctes r_1 et r_2 ; dans ce cas, les fonctions $x_1 : t \mapsto \mathrm{e}^{r_1 t}$ et $x_2 : t \mapsto \mathrm{e}^{r_2 t}$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :
      $$t \mapsto \lambda_1 \mathrm{e}^{r_1 t} + \lambda_2 \mathrm{e}^{r_2 t}\,,\ (\lambda_1,\lambda_2) \in \mathbb{C}^2\,.$$
    - Si $\Delta = 0$, (E_c) possède une racine double r ; dans ce cas, les fonctions $x_1 : t \mapsto \mathrm{e}^{rt}$ et $x_2 : t \mapsto t\mathrm{e}^{rt}$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :
      $$t \mapsto (\lambda_1 t + \lambda_2)\mathrm{e}^{rt}\,,\ (\lambda_1,\lambda_2) \in \mathbb{C}^2\,.$$
  - Si $\mathbb{K} = \mathbb{R}$
    - Si $\Delta \geqslant 0$, on a les mêmes résultats que ci-dessus, mais avec $(\lambda_1,\lambda_2) \in \mathbb{R}^2$.
    - Si $\Delta < 0$, l'équation (E_c) possède deux racines complexes non réelles conjuguées $\alpha \pm \mathrm{i}\beta$; alors les fonctions $x_1 : t \mapsto \mathrm{e}^{\alpha t}\cos\beta t$ et $x_2 : t \mapsto \mathrm{e}^{\alpha t}\sin\beta t$ sont deux solutions linéairement indépendantes de (H), et $\mathscr{S}_H$ est l'ensemble des applications de la forme :
      $$t \mapsto (\lambda_1 \cos\beta t + \lambda_2 \sin\beta t)\mathrm{e}^{\alpha t}\,,\ (\lambda_1,\lambda_2) \in \mathbb{R}^2\,.$$

- **Résolution de l'équation avec second membre**

  Il y a bien sûr toujours la méthode de variation des constantes, puisque l'on vient de trouver deux solutions linéairement indépendantes de l'équation homogène. Il s'agit de la méthode à préférer mais il y a un cas particulier à connaître.

  Si le second membre est de la forme $\mathrm{e}^{mt}P(t)$, où P est un polynôme et $m \in \mathbb{C}$, on cherchera une solution particulière de (L) sous la forme :

  $$x(t) = t^k \mathrm{e}^{mt} Q(t) \text{ où } \begin{cases} k = 0 & \text{si } m \text{ n'est pas racine de } (E_c) \\ k = 1 & \text{si } m \text{ est racine simple de } (E_c) \\ k = 2 & \text{si } m \text{ est racine double de } (E_c) \end{cases}$$

  avec Q polynôme de même degré que P.

Thème 18 - Calcul différentiel

On étudie ici des fonctions définies sur un $\mathbb{R}$-espace vectoriel normé E de dimension $p \in \mathbb{N}^*$, à valeurs dans un $\mathbb{R}$-espace vectoriel normé F de dimension $n \in \mathbb{N}^*$. Quitte à choisir des bases dans ces espaces vectoriels, on pourra se ramener à l'étude de fonctions de $\mathbb{R}^p$ dans $\mathbb{R}^n$.

Les espaces vectoriels considérés étant de dimensions finies, toutes les normes y sont équivalentes. On notera donc $\|\ \|$ une norme quelconque sur E ou F.

Les applications considérées seront toujours définies sur un *ouvert* de E ; cela assure que, si f est définie en $a \in U$, alors elle reste définie dans tout un voisinage de a.

[S18.1] Rappels

- **Applications coordonnées**
  Si f est une application de $U \subset \mathbb{R}^p$ dans $\mathbb{R}^n$, on a :
  $$\forall x \in U,\ f(x) = \big(f_1(x), \dots, f_n(x)\big)$$
  où les f_i sont des applications de U dans $\mathbb{R}$: ce sont les applications coordonnées de f.
  L'utilisation de ces applications coordonnées permet de ramener l'étude des fonction de $\mathbb{R}^p$ dans $\mathbb{R}^n$ à celle des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$.
- **Applications partielles**
  Soit $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$, et $a = (a_1, \dots, a_p) \in U$. Pour tout $j \in [\![1\,;p]\!]$, la j-ème application partielle de f en a est l'application :
  $$f_{a,j} : x \mapsto f(a_1, \dots, a_{j-1}, x, a_{j+1}, \dots, a_p)\,.$$
  Elle est définie sur un ouvert de $\mathbb{R}$ contenant a_j et à valeurs dans $\mathbb{R}^n$.
- **Continuité**
  Une application $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ d'un ouvert U de $\mathbb{R}^p$ à valeurs dans $\mathbb{R}^n$ est dite continue en $a \in U$ si :
  $$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in U,\ \|x - a\| < \alpha \implies \|f(x) - f(a)\| < \varepsilon$$
  ou, plus simplement :
  $$\lim_{x \to a} f(x) = f(a)\,.$$
- Une combinaison linéaire, une composée, un produit ou un quotient (pour des applications à valeurs dans $\mathbb{R}$) d'applications continues est une application continue là où elle est définie.
- Une application $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ est continue si et seulement si ses applications coordonnées le sont.
- Si $f : U \subset \mathbb{R}^p \to \mathbb{R}^n$ est continue en $a \in U$, alors toutes ses applications partielles $f_{a,j}$ en ce point sont continues.

  ✓ La réciproque de cette propriété est fausse : la continuité des seules applications partielles ne suffit pas à assurer la continuité en un point.

[S18.2] Différentiabilité

- **Différentielle**

  Soit f une application définie sur un ouvert U de E à valeurs dans F.

  On dit que f est différentiable en un point a de U s'il existe une application linéaire $\mathrm{d}f_a \in \mathscr{L}(E,F)$ telle que :

  $$\text{pour tout } h \in E \text{ tel que } a+h \in U,\ f(a+h) = f(a) + \mathrm{d}f_a(h) + o(\|h\|)\,.$$

  $\mathrm{d}f_a$ s'appelle la différentielle de f en a (ou application linéaire tangente à f en a). L'écriture ci-dessus s'appelle un développement limité d'ordre 1 de f en a.

- Si f est différentiable en a, alors f est continue en a.

  ✓ La réciproque de cette propriété est fausse !

- **Dérivée selon un vecteur**

  Soit f une application définie sur un ouvert U de E à valeurs dans F, et a un point de U. Soit $\overrightarrow{v}$ un vecteur non nul de E, et t un réel tel que $a + t\overrightarrow{v} \in U$.

  On dit que f admet une dérivée selon le vecteur $\overrightarrow{v}$ au point a si la fonction $t \mapsto f(a+t\overrightarrow{v})$ est dérivable en 0.

  Le vecteur dérivé correspondant est appelé dérivée de f suivant le vecteur $\overrightarrow{v}$, et est noté $\dfrac{\partial f}{\partial \overrightarrow{v}}(a)$ ou $D_{\overrightarrow{v}} f(a)$. Ainsi :

  $$\frac{\partial f}{\partial \overrightarrow{v}}(a) = D_{\overrightarrow{v}} f(a) = \lim_{t\to 0} \frac{f(a+t\overrightarrow{v}) - f(a)}{t}\,.$$

- **Théorème**

  Soit f une application définie sur un ouvert U de E à valeurs dans F, et a un point de U. Si f est différentiable en a, alors pour tout vecteur $\overrightarrow{v}$ de E, f admet une dérivée selon $\overrightarrow{v}$ et :

  $$D_{\overrightarrow{v}} f(a) = \mathrm{d}f_a(\overrightarrow{v})\,.$$

  ✓ La réciproque est fausse : une fonction peut admettre une dérivée en un point selon n'importe quel vecteur sans pour autant être différentiable en ce point.

[S18.3] Dérivées partielles

- Soit f une application définie sur un ouvert U de E à valeurs dans F, et a un point de U.

  Soit $\mathscr{B}_E = (\overrightarrow{e_1},\dots,\overrightarrow{e_p})$ une base de E ; on confondra alors tout vecteur $x \in E$ avec le p-uplet de ses coordonnées $(x_1,\dots,x_p)$, et on écrira indifféremment $f(x)$ ou $f(x_1,\dots,x_p)$

  On appelle dérivées partielles (premières) de f ses dérivées suivant les vecteurs $\overrightarrow{e_1},\dots,\overrightarrow{e_p}$ (si elles existent).

  La j^{e} dérivée partielle en a est notée $\partial_j f(a)$ ou $\dfrac{\partial f}{\partial x_j}(a)$ (c'est un vecteur appartenant à F). Ainsi :

  $$\forall\, j \in [\![1\,;p]\!],\ \frac{\partial f}{\partial x_j}(a) = D_{\overrightarrow{e_j}} f(a) = \lim_{t\to 0} \frac{f(a+t\overrightarrow{e_j}) - f(a)}{t}\,.$$

Si $a = (a_1, \ldots, a_p)$, c'est donc aussi la dérivée en a_j de la j^{e} application partielle $f_{a,j}$:

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t \to 0} \frac{f(a_1, \ldots, a_{j-1}, a_j + t, a_{j+1}, \ldots, a_p) - f(a_1, \ldots, a_p)}{t}.$$

- Si f est différentiable en a, alors f admet en ce point des dérivées partielles par rapport à chacune des variables et l'on a :

$$\forall\, h = (h_1, \ldots, h_p) \in E,\ \mathrm{d}f_a(h) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a).$$

[S18.4] Matrice jacobienne

- Soit f une application définie sur un ouvert U de E à valeurs dans F. On choisit dans E une base $\mathscr{B}_E$ et dans F une base $\mathscr{B}'_F$.

  On notera $f_1, \ldots, f_n$ les applications coordonnées de f dans la base $\mathscr{B}'_F$, c'est-à-dire que, pour tout $x \in U$, $f(x) = \big(f_1(x), \ldots, f_n(x)\big)$. Les f_i sont des applications de U dans $\mathbb{R}$.

  Soit $a \in U$. On suppose f différentiable en a. On appelle alors matrice jacobienne de f en a, notée $J_f(a)$, la matrice de la différentielle $\mathrm{d}f_a$ de f en a dans les bases $\mathscr{B}_E$ et $\mathscr{B}'_F$.

  $J_f(a)$ est donc la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ dont le terme d'indice $(i,j) \in [\![1\,;n]\!] \times [\![1\,;p]\!]$ est égal à $\dfrac{\partial f_i}{\partial x_j}(a)$:

$$J_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_p}(a) \\ \vdots & & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \dfrac{\partial f_n}{\partial x_2}(a) & \cdots & \dfrac{\partial f_n}{\partial x_p}(a) \end{pmatrix}.$$

- Lorsque $n = p$, on appelle jacobien de f en a le déterminant de sa matrice jacobienne.

  On le note, de façon abrégée : $\dfrac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}$.

[S18.5] Opérations sur les fonctions différentiables

- **Applications coordonnées**

  Soit f une application définie sur un ouvert U de E à valeurs dans F. On choisit dans F une base $\mathscr{B}'_F$, et on note $f_1, \ldots, f_n$ les applications coordonnées de f dans cette base.

  Alors f est différentiable en $a \in U$ si et seulement si pour tout $i \in [\![1\,;n]\!]$ f_i est différentiable en a ; dans ce cas, les applications coordonnées de $\mathrm{d}f_a$ dans $\mathscr{B}'_F$ sont les $\mathrm{d}(f_i)_a$.

Ce résultat permet de ramener l'étude de la différentiabilité à celle des fonctions à valeurs réelles.

- **Combinaison linéaire**

  Soient f et g deux applications d'un ouvert U de E à valeurs dans F, différentiables en $a \in U$.

  Pour tous réels λ et μ, l'application $\lambda f + \mu g$ est différentiable en a et :

  $$\mathrm{d}(\lambda f + \mu g)_a = \lambda\,\mathrm{d}f_a + \mu\,\mathrm{d}g_a\,,$$

  ou encore, des bases ayant été choisies dans E et F : $J_{\lambda f+\mu g}(a) = \lambda J_f(a) + \mu J_g(a)$.

- **Bilinéarité**

  Soient E, F_1, F_2, G des espaces vectoriels normés de dimensions finies, et $B\colon F_1 \times F_2 \to G$ une application bilinéaire.

  Si $f\colon U \to F_1$ et $g\colon U \to F_2$ sont différentiables en $a \in U$, alors $B(f,g)$ est différentiable en a et :

  $$\forall\, h \in E,\ \mathrm{d}B(f,g)_a(h) = B\big(\mathrm{d}f_a(h), g(a)\big) + B\big(f(a), \mathrm{d}g_a(h)\big)\,.$$

- **Composée**

  - Soit f une application définie sur un ouvert $U \subset E$, à valeurs dans un ouvert V de F, et g une application de V dans un espace vectoriel normé de dimension finie G.

    On suppose f différentiable en $a \in U$, et g différentiable en $b = f(a) \in V$. Alors l'application composée $g \circ f$ est différentiable en a et l'on a :

    $$\mathrm{d}(g \circ f)_a = \mathrm{d}g_{f(a)} \circ \mathrm{d}f_a$$

    (on rappelle que $\mathrm{d}f_a \in \mathscr{L}(E,F)$, $\mathrm{d}g_{f(a)} \in \mathscr{L}(F,G)$ et $\mathrm{d}(g \circ f)_a \in \mathscr{L}(E,G)$).

    En choisissant des bases dans E, F et G, on a donc la relation suivante pour les matrices jacobiennes dans ces bases :

    $$J_{g\circ f}(a) = J_g(f(a)) \times J_f(a)\,.$$

  - Soient $p = \dim E$, $q = \dim F$ et $n = \dim G$. En notant $h = g \circ f$, $(h_1, \ldots, h_n)$ ses applications coordonnées, $(g_1, \ldots, g_n)$ celles de g et $(f_1, \ldots, f_q)$ celles de f, puis en notant $(x_1, \ldots, x_p)$ les coordonnées d'un élément de E et $(y_1, \ldots, y_q)$ celles d'un élément de F (cela correspond à $y = f(x)$), la formule précédente se traduit par :

    $$\forall\, i \in [\![1\,;n]\!],\ \forall\, j \in [\![1\,;p]\!],\ \frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{q} \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)\,.$$

    C'est la *règle de dérivation en chaîne.*

  - **Cas particulier : dérivée le long d'un arc**

    Soient les applications $f\colon U \to F$, avec U ouvert de E, et $\gamma\colon I \to E$, avec I intervalle de $\mathbb{R}$, telles que $\gamma(I) \subset U$.

    Si γ est dérivable en t et si f est différentiable en $\gamma(t)$, alors $g = f \circ \gamma$ est dérivable en t et l'on a :

    $$g'(t) = (f \circ \gamma)'(t) = \mathrm{d}f_{\gamma(t)}\big(\gamma'(t)\big)\,.$$

En notant $(\gamma_1, \ldots, \gamma_p)$ les fonctions coordonnées de γ, cela s'écrit :

$$g'(t) = \sum_{j=1}^{p} \gamma_j'(t) \frac{\partial f}{\partial x_j}(\gamma_1(t), \ldots, \gamma_p(t)) .$$

Géométriquement et en simplifiant, cela signifie que si Γ est un arc paramétré régulier et si f est différentiable sur un ouvert contenant le support de cet arc, un vecteur directeur de la tangente à $f(\Gamma)$ au point de paramètre t est l'image par $\mathrm{d}f$ d'un vecteur directeur de la tangente à Γ au point de paramètre t (si ce vecteur est non nul).

[S18.6] Gradient

- **Représentation des formes linéaires d'un espace euclidien**
  Soit φ une forme linéaire sur un espace euclidien E.
  Alors il existe un et un seul vecteur $a \in E$ tel que :

$$\forall x \in E,\ \varphi(x) = \langle a | x \rangle .$$

  L'application qui à tout vecteur $a \in E$ associe la forme linéaire $\varphi_a : x \mapsto \langle a | x \rangle$ s'appelle l'isomorphisme canonique entre E et son dual E^*.

- Soit f une application définie sur un ouvert U de E euclidien, à valeurs dans $\mathbb{R}$, et différentiable en $a \in U$. La différentielle $\mathrm{d}f_a$ est donc une forme linéaire sur E.
  On appelle gradient de f en a , noté $\overrightarrow{\mathrm{grad}} f(a)$ ou $\nabla f(a)$ l'unique vecteur de E tel que :

$$\forall h \in E,\ \mathrm{d}f_a(h) = \langle \nabla f(a) | h \rangle .$$

  Si E est rapporté à une base orthonormale $\mathscr{B}_E$, les coordonnées de $\nabla f(a)$ dans $\mathscr{B}_E$ sont :

$$\left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_p}(a) \right) .$$

- Compte tenu de l'inégalité de Cauchy-Schwarz, la relation $\dfrac{\partial f}{\partial \overrightarrow{v}}(a) = \langle \nabla f(a) | \overrightarrow{v} \rangle$ montre que le vecteur gradient (lorsqu'il est non nul) donne la direction selon laquelle la dérivée de f est maximale.

[S18.7] Extrema

Soit f une application définie sur un ouvert U de E à valeurs dans $\mathbb{R}$. Soit $a \in U$.

- On dit que f admet un minimum (respectivement un maximum) global en a si :

$$\forall x \in U,\ f(x) \geqslant f(a) \quad (\text{respectivement } f(x) \leqslant f(a)) .$$

- On dit que f admet un minimum (respectivement un maximum) local en a si :

$$\exists V \in \mathcal{V}(a),\ \forall x \in V,\ f(x) \geqslant f(a) \quad (\text{respectivement } f(x) \leqslant f(a)) .$$

- Si f admet un extremum local en a et est différentiable en a, alors $\nabla(f)(a) = 0$ (c'est-à-dire que pour tout $j \in [\![1\,;p]\!]$, $\dfrac{\partial f}{\partial x_j}(a) = 0$).

  ✓ La réciproque de ce théorème est fausse : le gradient de f peut s'annuler en a (a s'appelle alors un point critique de f) sans qu'il s'agisse d'un extremum.

[S18.8] Fonctions de classe $\mathscr{C}^1$

- Une application $f : U \to F$ est dite de classe $\mathscr{C}^1$ si f est différentiable sur U et si sa différentielle $\mathrm{d}f : \begin{cases} U & \longrightarrow & \mathscr{L}(E,F) \\ a & \longmapsto & \mathrm{d}f_a \end{cases}$ est continue sur U.

- **Caractérisation**

  Soient U un ouvert de E et f une application de U dans F.

  f est de classe $\mathscr{C}^1$ sur U si et seulement si f admet en tout point de U des dérivées partielles par rapport à chacune des variables, et si ces dérivées partielles sont continues sur U.

  ✓ Ainsi, la seule existence des dérivées partielles ne garantit même pas la continuité, alors que l'hypothèse supplémentaire de la continuité de ces dérivées partielles garantit l'existence (et la continuité) de la différentielle.

- **Opérations sur les fonctions de classe $\mathscr{C}^1$**

  - Si f et g sont de classe $\mathscr{C}^1$ de U dans F, alors pour tout λ réel, $\lambda f + g$ est de classe $\mathscr{C}^1$ sur U.

    Il en résulte que l'ensemble $\mathscr{C}^1(U,F)$ des fonctions de classe $\mathscr{C}^1$ de U dans F est un $\mathbb{R}$-espace vectoriel.

  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}$, leur produit fg est de classe $\mathscr{C}^1$ sur U et l'on a :

    $$\forall\, a \in U,\ \mathrm{d}(fg)_a = f(a)\,\mathrm{d}g_a + g(a)\,\mathrm{d}f_a\,.$$

    L'ensemble $\mathscr{C}^1(U,\mathbb{R})$ est donc une $\mathbb{R}$-algèbre.

  - Si f et g sont de classe $\mathscr{C}^1$ de U dans $\mathbb{R}$, leur quotient $\dfrac{f}{g}$ est de classe $\mathscr{C}^1$ sur $U' = U \setminus \{x \in U \mid g(x) = 0\}$ (qui est bien un ouvert) et l'on a :

    $$\forall\, a \in U',\ \mathrm{d}\left(\frac{f}{g}\right)_a = \frac{1}{g(a)^2}\left(g(a)\,\mathrm{d}f_a - f(a)\,\mathrm{d}g_a\right)\,.$$

  - Si f et g sont de classe $\mathscr{C}^1$ de U dans F_1 et F_2 respectivement, si B est une application bilinéaire de $F_1 \times F_2$ dans G, alors $B(f,g)$ est de classe $\mathscr{C}^1$ de U dans G.

  - Si F est de classe $\mathscr{C}^1$ sur U à valeurs dans un ouvert V de F et si g est de classe $\mathscr{C}^1$ de V dans G, alors $g \circ f$ est de classe $\mathscr{C}^1$ sur U.

- **Caractérisation des fonctions constantes**

  Une application f de classe $\mathscr{C}^1$ sur un ouvert *connexe par arcs* U de E à valeurs dans F est constante sur U si et seulement si sa différentielle $\mathrm{d}f$ est nulle en tout point de U.

  Cela équivaut à dire que toutes les dérivées partielles de f sont nulles sur U.

[S18.9] Dérivées partielles d'ordre supérieur

- Soit f une fonction définie sur un ouvert U de E à valeurs dans F. E est rapporté à une base $\mathscr{B}_E = (e_1, \dots, e_p)$. Soient $(i,j) \in [\![1\,;p]\!]^2$.

  On suppose que f possède une j^{e} dérivée partielle $\partial_j f$ ou $\dfrac{\partial f}{\partial x_j}$. Cette dernière peut elle-même posséder une i^{e} dérivée partielle qui est alors notée :

$$\partial^2_{i,j} f = \partial_i(\partial_j f) \quad \text{ou} \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right),$$

  et est appelée *dérivée partielle seconde* de f par rapport aux variables x_i, x_j.

- On dit que $f\colon U \to \mathbb{R}^n$ est de classe $\mathscr{C}^2$ sur U si toutes ses dérivées partielles secondes existent et sont continues sur U.

- **Théorème de Schwarz**

  Si $f \in \mathscr{C}^2(U,\mathbb{R}^n)$ où U est un ouvert de $\mathbb{R}^p$, on a :

$$\forall\,(i,j) \in [\![1\,;p]\!]^2,\ \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}\cdot$$

- **Fonctions de classe $\mathscr{C}^k$**

  Soient U un ouvert de E et f une application de U dans F.

  Soit k un entier non nul. Si cela est possible, on peut définir la dérivée partielle d'ordre k par rapport aux variables successives $x_{i_1}, \dots, x_{i_k}$ par

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\left(\cdots\left(\frac{\partial f}{\partial x_{i_1}}\right)\right)\right).$$

  On dit que f est de classe $\mathscr{C}^k$ sur U si et seulement si f admet des dérivées partielles successives jusqu'à l'ordre k par rapport à toutes les variables à toutes les places possibles, et si ces dérivées partielles sont toutes continues sur U.

- **Théorème de Schwarz généralisé**

  Si f est de classe $\mathscr{C}^k$ avec $k \geqslant 2$ sur un ouvert U de E, alors, pour tout entier N de $[\![2\,;k]\!]$, tout N-uplet d'indices $(i_1, \dots, i_N) \in [\![1\,;p]\!]^N$ et toute permutation $\sigma \in \mathfrak{S}_N$, on a

$$\frac{\partial^N f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_N}} = \frac{\partial^N f}{\partial x_{i_{\sigma(1)}} \partial x_{i_{\sigma(2)}} \dots \partial x_{i_{\sigma(N)}}}\cdot$$

- **Opérations sur les fonctions de classe $\mathscr{C}^k$**

  Les résultats sont similaires à ceux obtenus pour les fonctions de classe $\mathscr{C}^1$ (si f et g sont de classe $\mathscr{C}^k$, il en est de même de $\lambda f + g$ etc...).

[S18.10] Application aux courbes planes

- Une courbe du plan définie par une équation cartésienne est un ensemble de points (x,y) du plan vérifiant une équation de la forme :
$$f(x,y) = 0$$
où f est de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.
- Un point $M(x_0,y_0)$ de la courbe $\mathcal{C}$ d'équation $f(x,y) = 0$ est dit régulier si ce n'est pas un point critique de f, c'est-à-dire si :
$$\nabla f(x_0,y_0) \neq 0\,.$$
- Si M_0 est un point régulier de $\mathcal{C}$, la courbe admet en ce point une tangente ; c'est la droite passant par M_0 et de vecteur normal $\nabla f(x_0,y_0)$, c'est-à-dire d'équation :
$$\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) = 0\,.$$

[S18.11] Application aux surfaces

- Une surface définie par une équation cartésienne est un ensemble de points (x,y,z) de l'espace vérifiant une équation de la forme :
$$f(x,y,z) = 0$$
où f est une application de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^3$ et à valeurs dans $\mathbb{R}$.
- Un point $M(x_0,y_0,z_0)$ de la surface $\mathcal{S}$ d'équation $f(x,y,z) = 0$ est dit régulier si ce n'est pas un point critique de f, c'est-à-dire si :
$$\nabla f(x_0,y_0,z_0) \neq 0\,.$$
- Si M_0 est un point régulier de $\mathcal{S}$, la surface admet en ce point un plan tangent ; c'est le plan passant par M_0 et de vecteur normal $\nabla f(x_0,y_0,z_0)$, c'est-à-dire d'équation :
$$\frac{\partial f}{\partial x}(x_0,y_0,z_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0,z_0)(y-y_0) + \frac{\partial f}{\partial z}(x_0,y_0,z_0)(z-z_0) = 0\,.$$
La droite passant par M_0 et de vecteur directeur $\nabla f(x_0,y_0,z_0)$ s'appelle la normale à $\mathcal{S}$ en M_0.
- On appelle courbe tracée sur la surface $\mathcal{S}$ d'équation $f(x,y,z) = 0$ tout arc paramétré (I,γ) où I est un intervalle de $\mathbb{R}$ et où $\gamma : I \to \mathbb{R}^3$ est une application de classe $\mathscr{C}^1$ à valeurs dans U telle que :
$$\forall\, t \in I,\ f(\gamma(t)) = 0\,,$$
c'est-à-dire que pour tout t le point $M(t)$ de l'arc γ appartient à $\mathcal{S}$.
- En dérivant la relation précédente à l'aide de la règle de la chaîne on obtient :
$$\forall\, t \in I,\ \langle \nabla f(\gamma(t)) \,|\, \gamma'(t) \rangle = 0\,,$$
donc en tout point régulier (de l'arc et de la surface), la tangente à l'arc est orthogonale au gradient et par suite appartient au plan tangent à la surface.

SAVOIR-FAIRE

■ Comment montrer qu'un magma est un groupe ?

SF1.1 Revenir à la définition

Vérifier une par une les propriétés d'un groupe est souvent la seule solution lorsque la loi considérée n'est pas une loi connue.

Exemple

Soit $G = [0\,;1[\,\cap\,\mathbb{Q}$. On définit sur G une loi $*$ par :

$$\forall\,(x,y) \in G^2,\ x * y = x + y - \lfloor x + y \rfloor\ .$$

1) Montrer que $(G,*)$ est un groupe abélien.

2) Montrer que tous les éléments de G sont d'ordre fini.

1) • *La loi $*$ est bien définie sur G^2 ; de plus, par définition de la partie entière, on a bien $x * y \in [0\,;1[$ pour tous x, $y \in G$. Enfin, si x et y sont rationnels, $x * y$ est encore un rationnel.*
La loi $*$ est donc une loi de composition interne dans G.

• *L'associativité et la commutativité de $*$ sont immédiates : elles découlent des propriétés de l'addition dans $\mathbb{Q}$.*
• *Si $x \in G$, $\lfloor x \rfloor = 0$ donc $x * 0 = x$; on en déduit que 0 est l'élément neutre de G pour $*$.*

✓ *Nous avons seulement vérifié que 0 est neutre à droite, mais cela suffit puisque la loi $*$ est commutative.*
De plus, l'élément neutre étant unique, il est inutile de résoudre l'équation générale $e * x = x$: puisque 0 vérifie cette équation pour tout x, c'est nécessairement l'élément neutre.

• *Si $x \in G$ et $x \neq 0$, on a $1 - x \in G$ et il est facile de vérifier que $x * (1 - x) = 0$. La loi $*$ étant commutative, cela signifie que $1 - x$ est le symétrique de x. De plus, 0 est son propre symétrique, donc tout élément de G possède un symétrique dans G pour la loi $*$.*

✓ *Là encore, pour trouver le symétrique d'un élément $x \in G$, nul besoin de résoudre l'équation $x * x' = x' * x = e_G$; si l'on devine la valeur de x', c'est forcément le symétrique de x par unicité.*

• *En conclusion, $(G,*)$ est bien un groupe abélien.*

2) *Pour $x \in G$ et $n \in \mathbb{N}$, on vérifie par récurrence sur n (en utilisant le fait que la fonction partie entière est 1-périodique), que l'itéré n^e de x est : $x^{(n)} = nx - \lfloor nx \rfloor$.*
Soit alors $x = \dfrac{p}{q} \in G \setminus \{0\}$, avec p, $q \in \mathbb{N}^*$, $p \wedge q = 1$ et $p < q$.
L'itéré q^e de x est $x^{(q)} = qx - \lfloor qx \rfloor = p - p = 0 = e_G$, donc x est d'ordre fini.

Plus précisément, x est d'ordre exactement q puisque :

$$x^{(n)} = 0 \iff nx = \lfloor nx \rfloor \iff nx \in \mathbb{N} \iff \frac{np}{q} \in \mathbb{N} \iff q \mid n\,,$$

d'après le théorème de Gauss.

✓ Nous avons ici un exemple de groupe de cardinal infini dont tous les éléments sont d'ordre fini.

Revenir à la définition peut être aussi la seule solution dans le cadre d'exercices plus théoriques.

Exemple

Soit E un ensemble fini muni d'une loi de composition interne $*$, associative et pour laquelle tous les éléments sont réguliers.

Montrer que $(E,*)$ est un groupe.

- *Soit $x \in E$. L'application $\gamma_x : \begin{cases} E & \longrightarrow & E \\ y & \longmapsto & x*y \end{cases}$ est injective par hypothèse ; E étant de cardinal fini, elle est bijective. Il existe donc un élément $e_x \in E$ tel que $x * e_x = x$.*
  De même, en considérant l'application $\delta_x : y \mapsto y * x$, on obtient qu'il existe e'_x tel que $e'_x * x = x$.
  On a alors, en utilisant l'associativité de la loi :
  $$x * x = x * (e'_x * x) = (x * e'_x) * x\,,$$
  donc en utilisant la régularité de x à droite, on obtient $x = x*e'_x$. L'égalité $x * e'_x = x * e_x$ implique alors $e_x = e'_x$.
  On a donc montré pour tout $x \in E$ l'existence d'un élément e_x tel que $x * e_x = e_x * x = x$.
- *Pour x, $y \in E$ on a alors, toujours en utilisant l'associativité :*
  $$(x * e_x) * y = x * y = x * (e_y * y) = (x * e_y) * y\,,$$
  d'où en utilisant la régularité à droite de y, $x*e_x = x*e_y$, puis en utilisant la régularité à gauche de x : $e_x = e_y$.
  On a ainsi montré que la loi $*$ possède un élément neutre e.
- *Pour tout $x \in E$, il existe x' et x'' dans E tels que $\gamma_x(x) = e$ et $\delta_x(x) = e$, c'est-à-dire tels que :*
  $$x * x' = x'' * x = e\,.$$
  Alors $x'' = x''*(x*x') = (x''*x)*x' = x'$, et on a ainsi prouvé l'existence d'un symétrique pour tout élément de E, ce qui achève la démonstration.

SF1.2 Montrer qu'il s'agit d'un sous-groupe d'un groupe connu

Il s'agit de la solution la plus courante, et généralement la plus rapide.

Exemple

Soit $n \in \mathbb{N}^*$. Montrer que l'ensemble G des matrices carrées d'ordre n, triangulaires supérieures à éléments diagonaux strictement positifs, est un groupe pour la multiplication des matrices.

Montrons que G est un sous-groupe du groupe linéaire $(\mathrm{GL}_n(\mathbb{R}),\times)$.

- *Déjà, G est bien inclus dans $\mathrm{GL}_n(\mathbb{R})$ puisque si $A \in G$, on a $\det A > 0$.*
- *L'élément neutre I_n de $\mathrm{GL}_n(\mathbb{R})$ appartient bien à G.*
- *Soient A, $B \in G$, $A = (a_{ij})$ et $B = (b_{ij})$. On sait que le produit AB est encore une matrice triangulaire supérieure et que ses éléments diagonaux sont les produits $a_{ii}b_{ii}$; ils sont donc strictement positifs, d'où $AB \in G$.*
- *Soit $A \in G$. On sait alors que son inverse est triangulaire supérieure, et que ses coefficients diagonaux sont les inverses des coefficients diagonaux de A ; ils sont donc strictement positifs, et ainsi $A^{-1} \in G$.*

D'après la caractérisation du cours, G est bien un sous-groupe de $(\mathrm{GL}_n(\mathbb{R}),\times)$.

SF1.3 Utiliser un morphisme

Il s'agit d'utiliser le fait que l'image directe ou l'image réciproque d'un sous-groupe par un morphisme est un sous-groupe. Nous avons d'ailleurs déjà utilisé plusieurs fois cette propriété dans le cours (groupe alterné, groupe spécial orthogonal ...)

Exemple

On définit sur $G =]-1\,;1[$ une loi $*$ par :

$$\forall\,(x,y) \in G^2,\ x * y = \frac{x+y}{1+xy}\,.$$

Montrer que $(G,*)$ est un groupe abélien, et que la fonction th réalise un isomorphisme de groupes de $(\mathbb{R},+)$ sur $(G,*)$.

- *Démontrons d'abord que $(G,*)$ est un magma, c'est-à-dire que la loi $*$ est une loi de composition interne dans G.*
  - *La loi $*$ est bien définie sur $G \times G$, puisque si x, $y \in]-1\,;1[$, on a $1+xy > 0$.*
  - *Si x, $y \in G$, alors :*
  $$\frac{x+y}{1+xy} - 1 = \frac{x+y-1-xy}{1+xy} = \frac{(x-1)(1-y)}{1+xy} < 0\,,$$
  donc $x*y < 1$. On montre de même que $x*y > -1$, et ainsi $x*y \in G$.
- *Montrons que la fonction* th *est un isomorphisme du magma $(\mathbb{R},+)$ sur le magma $(G,*)$.*
  - *On sait déjà que la fonction* th *réalise une bijection de $\mathbb{R}$ sur $]-1\,;1[$.*
  - *Pour tout $(x,y) \in \mathbb{R}^2$, on a :*
  $$\mathrm{th}(x+y) = \frac{\mathrm{th}\,x + \mathrm{th}\,y}{1 + \mathrm{th}\,x.\,\mathrm{th}\,y} = \mathrm{th}\,x * \mathrm{th}\,y\,,$$
  donc th *réalise un morphisme de $(\mathbb{R},+)$ sur $(G,*)$.*

  ✓ *Si vous ne connaissez pas la formule d'addition pour la fonction* th, *vous pouvez la retrouver facilement en écrivant* $\mathrm{th}\,x = \dfrac{\mathrm{e}^{2x}-1}{\mathrm{e}^{2x}+1}$.
- *Par transport de structure, l'image du groupe abélien $(\mathbb{R},+)$ par cet isomorphisme est un groupe abélien.*

■ Comment trouver l'ordre d'un élément dans un groupe ?

SF1.4 Utiliser l'exposant

Dire qu'un élément a d'un groupe G est d'ordre d signifie que d est le plus petit entier naturel non nul tel que $a^d = e_G$. De façon plus précise :

si n est tel que $a^n = e_G$, alors d divise n.

Exemple

Soit n un entier $\geqslant 2$, et $a \in [\![1\,;n-1]\!]$.

Démontrer que l'ordre de $\overline{a}$ dans le groupe $(\mathbb{Z}/n\mathbb{Z},+)$ est égal à $\dfrac{a \vee n}{a}$.

L'ordre de $\overline{a}$ est le plus petit entier k tel que $k\overline{a} = \overline{0}$.

✓ *La loi étant ici l'addition, l'itéré k^{e} d'un élément x est kx et non x^k !*

Or $k\overline{a} = \overline{0}$ si et seulement si ka est multiple de n, ce qui équivaut à ka multiple commun à a et n, donc multiple de $a \vee n$, soit encore à : k multiple de $\dfrac{a \vee n}{a}$.

Le plus petit entier k cherché est donc bien égal à $\dfrac{a \vee n}{a}$ (ou aussi à $\dfrac{n}{a \wedge n}$).

Exemple

Soient a et b deux éléments d'un groupe G, d'ordres p et q respectivement.

1) On suppose ici que a et b commutent et que p et q sont premiers entre eux. Montrer que ab est d'ordre pq.

2) On suppose ici G abélien.

a) Montrer que si d est un diviseur de p, il existe dans G un élément d'ordre d.

b) Montrer qu'il existe dans G un élément d'ordre $p \vee q$.

3) Les deux questions suivantes montrent l'importance des hypothèses faites.

a) On ne suppose plus $p \wedge q = 1$. Montrer que ab n'est pas nécessairement d'ordre pq, ni d'ordre $p \vee q$.

b) Dans $G = (\mathrm{GL}_2(\mathbb{R}),\times)$, soient $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$.
Montrer que A et B sont d'ordres finis premiers entre eux, mais que AB est d'ordre infini.

1) • *Puisque a et b commutent, on a*

$$(ab)^{pq} = a^{pq}b^{pq} = (a^p)^q(b^q)^p = e_G\,,$$

donc l'ordre de ab divise pq.

• *Réciproquement, soit $k \in \mathbb{N}^*$ tel que $(ab)^k = e_G$. On a alors :*

$$a^{kq} = a^{kq}\underbrace{b^{kq}}_{=(b^q)^k=e_G} = (ab)^{kq} = e_G\,,$$

donc l'ordre p de a divise kq, donc divise k d'après le théorème de Gauss.

De la même façon, q divise k. p et q étant premiers entre eux, le produit pq divise k.

• *On en conclut que l'ordre de ab est exactement égal à pq.*

2) a) *Soit d un diviseur de p, et $x = a^{p/d}$.*

– *On a alors $x^d = a^p = e_G$, donc l'ordre de x divise d.*

– *D'autre part, si $k \in \mathbb{N}^*$ est tel que $x^k = e_G$, alors $a^{kp/d} = e_G$ donc $\frac{kp}{d}$ est un multiple de p. Cela implique que $\frac{k}{d}$ est entier, c'est-à-dire d divise k.*

Il en résulte que x est exactement d'ordre d.

b) *On décompose p et q en facteurs premiers, sous la forme :*

$$p = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \quad \text{et} \quad q = p_1^{\beta_1} \cdots p_r^{\beta_r}$$

avec p_i premiers et α_i, $\beta_i \in \mathbb{N}$ (certains exposants peuvent être nuls).

On sait qu'alors le ppcm de p et q vaut : $p \vee q = \prod_{i=1}^{r} p_i^{\max(\alpha_i,\beta_i)}$.

Pour tout $i \in [\![1\,;r]\!]$, il existe dans G un élément x_i d'ordre $p_i^{\max(\alpha_i,\beta_i)}$: pour l'obtenir, on applique la question précédente avec a et $d = p_i^{\alpha_i}$ si $\alpha_i \geqslant \beta_i$, ou bien avec b et $d = p_i^{\beta_i}$ si $\beta_i \geqslant \alpha_i$.

G étant abélien, les éléments x_i commutent entre eux, et le produit $x_1 \cdots x_r$ est d'ordre le produit des ordres des x_i d'après la question **1)** *et une récurrence, c'est-à-dire d'ordre $p \vee q$.*

3) a) *Soit a un élément d'ordre $n \geqslant 2$ d'un groupe G, et $b = a^{-1}$. Alors a et b sont d'ordre n, mais leur produit $ab = e_G$ est d'ordre 1 !*

b) *On vérifie que $A^4 = \mathrm{I}_2$ (A est la matrice d'une rotation d'angle $\frac{\pi}{2}$), et que $B^3 = \mathrm{I}_4$. Ce sont donc des éléments d'ordre 4 et 3 respectivement du groupe $(\mathrm{GL}_2(\mathbb{R}),\times)$.*

Cependant, $AB = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ donc par récurrence on a pour tout entier $n \in \mathbb{N}^*$: $(AB)^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

On n'aura donc jamais $(AB)^n = \mathrm{I}_2$ pour $n \in \mathbb{N}^*$: AB est d'ordre infini (ici le résultat de la question **1)** *ne peut s'appliquer car $AB \neq BA$).*

SF1.5 Utiliser le théorème de Lagrange et ses conséquences

Ce théorème est presque incontournable dès que l'on étudie un groupe fini.

Exemple

Soient p et q deux nombres premiers impairs tels que q divise $2^p - 1$.
Montrer que $q \equiv 1 \pmod{2p}$.

On se place dans le groupe $\left((\mathbb{Z}/q\mathbb{Z})^*,\times\right)$ (c'est bien un groupe puisque, q étant premier, $\mathbb{Z}/q\mathbb{Z}$ est un corps).

L'hypothèse s'écrit alors : $\overline{2}^p = \overline{1}$, c'est-à-dire que l'ordre de $\overline{2}$ divise p, et puisque $\overline{2} \neq \overline{1}$, cet ordre est égal à p puisque p est premier.
D'après le théorème de Lagrange, p divise $q-1$. De plus $q-1$ est pair, donc on a aussi : $2p$ divise $q-1$, ce qui est le résultat voulu.

Exemple

Soit G un groupe abélien d'ordre pq, où p et q sont deux nombres premiers distincts.

1) Montrer que G est cyclique.

2) Montrer par un exemple que ce résultat peut tomber en défaut si l'on ne suppose pas G abélien.

1) *D'après le théorème de Lagrange, l'ordre de tout élément de* $G \setminus \{e_G\}$ *divise* pq*, donc est égal à* p*,* q *ou* pq*.*

• *Supposons que* G *admette un élément* x *d'ordre* p *et un élément* y *d'ordre* q*. On a alors :*

$$(xy)^p = x^p y^p = y^p \neq e_G \quad \text{et} \quad (xy)^q = x^q y^q = x^q \neq e_G\,,$$

donc l'élément xy *(qui est différent de* e_G *car sinon* $y = x^{-1}$ *aurait même ordre que* x*) est d'ordre* pq*.*

• *Supposons par l'absurde que* G *n'est pas cyclique, c'est-à-dire ne possède pas d'élément d'ordre* pq*. D'après ce qui précède, tous les éléments de* G *différents de* e_G *seraient alors soit tous d'ordre* p*, soit tous d'ordre* q*.*
Si tous les éléments de $G \setminus \{e_G\}$ *étaient d'ordre* p *(par exemple), soit* x *un tel élément et* H *le sous-groupe de* G *engendré par* x*. Soit* $y \notin H$*. Puisque* G *est abélien, le groupe engendré par* xy *est égal à :*

$$\mathrm{gr}(xy) = \left\{x^n y^m \;\middle|\; (n,m) \in \mathbb{Z}^2\right\}\,,$$

et puisque x *et* y *sont d'ordres* p *:*

$$\mathrm{gr}(xy) = \{x^n y^m \mid n \in [\![0\,;p-1]\!],\ m \in [\![0\,;p-1]\!]\}\,.$$

Montrons que ce groupe possède alors exactement p^2 *éléments ; en effet, si* $(n,n',m,m') \in [\![0\,;p-1]\!]^4$ *sont tels que* $x^n y^m = x^{n'} y^{m'}$*, alors* $x^{n-n'} = y^{m'-m}$*. Si l'on avait* $n \neq n'$ *(et donc* $m \neq m'$*), on aurait* $y^{m'-m} \in H \setminus \{e_G\}$*. Mais le groupe engendré par* y *est d'ordre* p *premier, donc* $y^{m'-m}$ *en est un générateur, et l'on aurait donc* $y \in H$*, ce qui est exclu.*
Ainsi, lorsque $(n,n',m,m') \in [\![0\,;p-1]\!]^4$*,* $(n,m) \neq (n',m')$ *implique* $x^n y^m \neq x^{n'} y^{m'}$*, donc il y a exactement* p^2 *éléments dans* $\mathrm{gr}(xy)$*, ce qui entraîne* $p^2 | pq$ *d'après le théorème de Lagrange, ce qui est absurde.*

• *On a ainsi prouvé par l'absurde que* G *est cyclique.*

2) *Pour* $G = \mathfrak{S}_3$ *(groupe symétrique d'ordre 3), on a* $\mathrm{Card}\, G = 6 = 2 \times 3$*, cependant* G *n'est pas cyclique.*

SF1.6 Utiliser un morphisme

Si $(G,*)$ est un groupe noté multiplicativement et si $a \in G$, l'application $\varphi : \begin{cases} \mathbb{Z} & \longrightarrow & G \\ k & \longmapsto & a^k \end{cases}$ est un morphisme du groupe $(\mathbb{Z},+)$ dans $(G,*)$.

Dire que a est d'ordre fini n équivaut à dire que le noyau de ce morphisme est égal à $n\mathbb{Z}$.

Exemple

Soit G un groupe cyclique de cardinal n.

Montrer que si d est un diviseur de n, alors G possède un unique sous-groupe de cardinal d et que celui-ci est cyclique.

On notera a un générateur de G et $\varphi\colon a \mapsto a^k$ le morphisme de $\mathbb{Z}$ dans G associé.
Soit H un sous-groupe de G.
$\varphi^{-1}(H)$ est un sous-groupe de $\mathbb{Z}$ contenant le noyau de φ c'est-à-dire $n\mathbb{Z}$. Il est donc de la forme $q\mathbb{Z}$ avec $q \in \mathbb{N}$, et puisque $n\mathbb{Z} \subset q\mathbb{Z}$, q divise n.
Par définition, un élément a^k de G appartient à H si et seulement si k appartient à $q\mathbb{Z}$. Donc $H = \{a^{qr} \mid r \in \mathbb{Z}\} = \{(a^q)^r \mid r \in \mathbb{Z}\}$ est exactement le sous-groupe de G engendré par a^q.
Les sous-groupes de G sont donc ceux de la forme $H = \mathrm{gr}(a^q)$ où q divise n. Notons $n = qd$. Si $k \in \mathbb{Z}$ on a alors :

$$(a^q)^k = e_G \iff n \mid qk \iff d \mid k\,,$$

donc l'ordre de a^q est égal à d, c'est-à-dire que le cardinal de H est égal à d. Il y a donc bien un seul sous-groupe d'ordre d, c'est celui engendré par a^q où $q = \dfrac{n}{d}$.

■ Comment montrer qu'un ensemble est un anneau, un corps, un idéal ?

SF1.7 Montrer qu'il s'agit d'une sous-structure

Heureusement, vous n'aurez pratiquement jamais à vérifier une par une toutes les hypothèses de la définition d'un anneau ou d'un corps : on montre le plus souvent qu'il s'agit d'un sous-anneau ou d'un sous-corps d'un anneau ou d'un corps connu.

Exemple

1) Soit $\mathbb{K}$ un sous-corps de $\mathbb{C}$, et n un entier naturel tel que $\sqrt{n} \notin \mathbb{K}$. On note alors :

$$\mathbb{K}[\sqrt{n}] = \left\{a + b\sqrt{n} \;\middle|\; (a,b) \in \mathbb{K}^2\right\}.$$

Montrer que $\mathbb{K}[\sqrt{n}]$ est un sous-corps de $\mathbb{C}$.

2) Montrer que $\mathbb{L} = \left\{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \;\middle|\; (a,b,c,d) \in \mathbb{Q}^4\right\}$, muni de l'addition et de la multiplication usuelles, est un corps.

1) *On vérifie toutes les caractérisations d'un sous-corps.*

- $1 = 1 + 0.\sqrt{n}$ *appartient à* $\mathbb{K}[\sqrt{n}]$.
- *Si* $x = a + b\sqrt{n}$ *et* $y = a' + b'\sqrt{n}$ *sont des éléments de* $\mathbb{K}[\sqrt{n}]$, *il est facile de vérifier que* $x - y$ *et* xy *sont encore des éléments de* $\mathbb{K}[\sqrt{n}]$.
- *Pour l'inverse, on remarque d'abord que, pour* $(a,b) \in \mathbb{K}^2$, $a + b\sqrt{n} = 0 \implies a = b = 0$ *puisque* $\sqrt{n} \notin \mathbb{K}$.
  Donc si $(a,b) \neq (0,0)$, *c'est-à-dire si* $a + b\sqrt{n} \neq 0$ *et* $a - b\sqrt{n} \neq 0$,

$$\frac{1}{a + b\sqrt{n}} = \frac{a - b\sqrt{n}}{a^2 - nb^2} = \frac{a}{a^2 - nb^2} - \frac{b}{a^2 - nb^2}\sqrt{n} \in \mathbb{K}[\sqrt{n}]\,.$$

2) *D'après ce qui précède,* $\mathbb{Q}[\sqrt{2}]$ *est un sous-corps de* $\mathbb{C}$ *(*$\sqrt{2} \notin \mathbb{Q}$*).*
Vérifions que $\sqrt{3}$ *n'appartient pas à* $\mathbb{Q}[\sqrt{2}]$ *: si tel était le cas, il existerait* $(a,b) \in \mathbb{Q}^2$ *tel que* $\sqrt{3} = a + b\sqrt{2}$*. On en déduit, en élevant au carré,* $3 - a^2 - 2b^2 = 2ab\sqrt{2}$*, donc* $ab = 0$ *sinon* $\sqrt{2}$ *serait rationnel ;* $a = 0$ *implique* $\sqrt{\frac{3}{2}}$ *rationnel, et* $b = 0$ *implique* $\sqrt{3}$ *rationnel : on aboutit à une contradiction.*
Par conséquent, d'après la question précédente, $\mathbb{L} = \mathbb{Q}[\sqrt{2}][\sqrt{3}]$ *est aussi un sous-corps de* $\mathbb{C}$*.*

SF1.8 Revenir à la définition

Pour montrer qu'une partie I d'un anneau commutatif A est un idéal de A, on revient le plus souvent à la caractérisation du cours.

Exemple

Soit D l'ensemble des nombres décimaux :

$$D = \{x \in \mathbb{Q} \mid \exists n \in \mathbb{N},\ 10^n x \in \mathbb{Z}\}\,.$$

Montrer que D est un anneau et que tous les idéaux de D sont principaux.

- *Montrons que* D *est un sous-anneau de* $\mathbb{Q}$*.*
  - $D \subset \mathbb{Q}$ *par définition, et* $0 \in D$*, donc* D *est non vide.*
  - *Soient* $x,\ y \in D$ *: il existe des entiers* $n,\ m$ *tels que* $10^n x \in \mathbb{Z}$ *et* $10^m y \in \mathbb{Z}$*. Alors :*
    - $10^{n+m}(x-y) = 10^m(10^n x) - 10^n(10^m y) \in \mathbb{Z}$*, donc* $x - y \in D$ *;*
    - $10^{n+m}xy = (10^n x)(10^m y) \in \mathbb{Z}$*, donc* $xy \in D$*.*

  D est donc bien un sous-anneau de $(\mathbb{Q}, +, \times)$*.*
- *Soit* I *un idéal de* D*.*
  - *Si* $I = \{0\}$*,* I *est principal.*
  - *Sinon,* $I \cap \mathbb{N}^*$ *est non vide puisque, si* $x \in I$*,* $\pm 10^n x \in I$ *pour tout* $n \in \mathbb{N}^*$ *par définition d'un idéal. Soit* a *le plus petit élément de* $I \cap \mathbb{N}^*$*. C'est aussi le plus petit élément strictement positif de* $I \cap \mathbb{Z}$*, et puisque* $(I \cap \mathbb{Z}, +)$ *est un sous-groupe de* $(\mathbb{Z}, +)$ *(comme intersection de deux sous-groupes de* $(\mathbb{Q}, +)$*), on a d'après le cours :* $I \cap \mathbb{Z} = a\mathbb{Z}$*.*

    Montrons que $I = aD$*.*
    - *On a évidemment* $aD \subset I$ *puisque* $a \in D$*.*
    - *Soit* $x \neq 0 \in I$*. Par définition de* D*, il existe* $p \in \mathbb{N}$ *et* $n \in \mathbb{N}$ *tels que* $x = \pm\frac{p}{10^n}$*. Puisque* 10^n *appartient à* D*,* $\pm 10^n x$ *appartient à* I *par définition d'un idéal, donc* $p \in I$*. Ainsi* $p \in I \cap \mathbb{Z} = a\mathbb{Z}$ *donc il existe* $k \in \mathbb{Z}$ *tel que* $p = ka$*.*
      On a donc $x = \pm\frac{k}{10^n}\, a$ *donc* $x \in aD$

  On a donc prouvé que tout idéal de D *est de la forme* aD*, c'est-à-dire est principal.*

SF1.9 Utiliser un morphisme

Exemple

Soit $a \in \mathbb{R}$. Dans l'anneau $A = (\mathcal{A}(\mathbb{R},\mathbb{R}), +, \times)$ des applications de $\mathbb{R}$ dans $\mathbb{R}$, on considère l'ensemble $I = \{f \in A \mid f(a) = 0\}$.

Montrer qu'il s'agit d'un idéal de A, et que cet idéal est maximal (c'est-à-dire que si un idéal J contient strictement I, alors $J = A$).

- *Par définition des lois d'addition et de multiplication des applications de* $\mathbb{R}$ *dans* $\mathbb{R}$, *l'application* $\varphi : \begin{cases} A & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f(a) \end{cases}$ *est un morphisme d'anneaux.*
  I en est le noyau, c'est donc un idéal de A d'après un résultat du cours.
- *Soit J un idéal de A contenant strictement I. Il existe donc $g \in J$ telle que $g(a) \neq 0$. Soit $f : x \mapsto x - a$. Alors f appartient à I donc à J, et puisque J est un idéal, la fonction $h = f^2 + g^2$ appartient à J.*
  Mais h ne s'annule pas sur $\mathbb{R}$, donc est un élément inversible de l'anneau A pour la loi $\times$. Puisque J est un idéal qui contient un élément inversible de l'anneau A, il est égal à A. Cela démontre le résultat demandé.

Exemple

Dans $\mathbb{R}$, on définit les deux lois :

$$\forall\,(a,b) \in \mathbb{R}^2,\ \begin{cases} a \star b = a + b - 1 \\ a \circ b = ab - (a+b) + 2\cdot \end{cases}$$

Quelle est la structure de $(\mathbb{R}, \star, \circ)$?

Il est facile de vérifier que $\star$ et $\circ$ sont des lois de composition internes sur $\mathbb{R}^2$. Soit $f : x \in \mathbb{R} \mapsto x - 1$. f est une bijection de $\mathbb{R}$ sur $\mathbb{R}$ et, pour tout $(a,b) \in \mathbb{R}^2$:

$$f(a \star b) = a + b - 2 = f(a) + f(b)\,,$$
$$f(a \circ b) = ab - a - b + 1 = (a-1)(b-1) = f(a)f(b)\,.$$

f est donc un morphisme pour les lois $\star$ et $+$ d'une part, et les lois $\circ$ et $\times$ d'autre part. Par transport de structure, l'image du corps $(\mathbb{R}, +, \times)$, c'est-à-dire $(\mathbb{R}, \star, \circ)$, est un corps.

■ Comment utiliser les congruences dans $\mathbb{Z}$?

SF1.10 Utiliser la compatibilité avec les lois + et ×

Il s'agit ici d'utiliser le fait que l'on peut additionner et multiplier les congruences modulo un même entier.

Exemple

Déterminer pour quelle(s) valeur(s) de $n \in \mathbb{N}$ l'entier : $A_n = 851^{3n} + 851^{2n} + 851^n$ est divisible par 7.

$851 = 121 \times 7 + 4$, *donc* $851 \equiv 4 \pmod 7$. *On en déduit :*

$$A_n \equiv 4^{3n} + 4^{2n} + 4^n \pmod 7 .$$

Puisque $4^2 \equiv 2 \pmod 7$ *puis* $4^3 \equiv 8 \equiv 1 \pmod 7$, *on a plus simplement :*

$$A_n \equiv 1 + 2^n + 4^n \pmod 7 ,$$

puis en effectuant la division euclidienne de n *par* 3 *: si* $n = 3q + r$ *avec* $r \in \{0,1,2\}$, *on a :*

$$A_n \equiv 1 + (2^3)^q 2^r + (4^3)^q 4^r \equiv 1 + 2^r + 4^r \pmod 7 .$$

Finalement :

$$\text{si } r = 0,\ A_n \equiv 3 \pmod 7; \quad \text{si } r = 1,\ A_n \equiv 7 \equiv 0 \pmod 7;$$
$$\text{si } r = 2,\ A_n \equiv 21 \equiv 0 \pmod 7.$$

Ainsi, A_n *est divisible par* 7 *si et seulement si* n *n'est pas multiple de* 3.

SF1.11 Faire les calculs dans $\mathbb{Z}/n\mathbb{Z}$

$\mathbb{Z}/n\mathbb{Z}$ est un anneau commutatif, on peut donc utiliser les règles de calculs habituelles dans un anneau. Lorsque n est premier, on peut aussi utiliser le fait que c'est un corps, c'est-à-dire que tout élément non nul y est inversible.

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{Z})$ telle que les termes diagonaux de A sont des nombres impairs et tous les autres sont des nombres pairs.
Montrer que A est inversible dans $\mathcal{M}_n(\mathbb{R})$.

Pour $x \in \mathbb{Z}$, *notons* $\overline{x}$ *sa classe d'équivalence modulo* 2. *Par définition des lois dans* $\mathbb{Z}/2\mathbb{Z}$, *on a, en notant* $A = (a_{ij})$ *:*

$$\overline{\det A} = \overline{\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \overline{a_{\sigma(i)i}} = \det \overline{A} ,$$

où $\overline{A}$ *est la matrice* $(\overline{a_{ij}})$.
Or compte tenu des hypothèses, $\overline{A} = \overline{I_n}$ *donc* $\det(\overline{A}) = \overline{1}$.
Cela montre que $\det A$ *est un entier impair, donc il est non nul et* A *est inversible.*

Exemple

Déterminer les $(x,y) \in \mathbb{Z}^2$ tels que $\begin{cases} 6x + 7y \equiv 30 \pmod{37} \\ 3x - 7y \equiv 0 \pmod{37} . \end{cases}$

Dans $\mathbb{Z}/37\mathbb{Z}$, *qui est un corps puisque* 37 *est premier, le système s'écrit :*

$$\begin{cases} \overline{6}\,\overline{x} + \overline{7}\,\overline{y} = \overline{30} \\ \overline{3}\,\overline{x} - \overline{7}\,\overline{y} = \overline{0} \end{cases} .$$

L'addition des deux équations implique $\overline{9}\,\overline{x} = \overline{30}$. *Mais l'inverse de* $\overline{9}$ *dans* $\mathbb{Z}/37\mathbb{Z}$ *est égal à* $-\overline{4}$ *(*$-4 \times 9 = -36 \equiv 1 \pmod{37}$*). Donc* $\overline{x} = -\overline{4} \times \overline{30} = -\overline{9} = \overline{28}$. *On a ensuite* $\overline{7}\,\overline{y} = \overline{3} \times \overline{28} = \overline{10}$. *L'inverse de* $\overline{7}$ *est* $\overline{16}$ *donc* $\overline{y} = \overline{16} \times \overline{10} = \overline{12}$.
L'ensemble des solutions dans $\mathbb{Z}^2$ *du système proposé est donc l'ensemble des couples :*

$$(x,y) = (28+37k, 12+37k') \text{ avec } (k,k') \in \mathbb{Z}^2 .$$

✓ *Puisque $\mathbb{Z}/37\mathbb{Z}$ est un corps, on pouvait aussi résoudre le système en utilisant les formules de Cramer ; en effet, le déterminant du système est $\begin{vmatrix} \overline{6} & \overline{7} \\ \overline{3} & -\overline{7} \end{vmatrix} = -\overline{63} = \overline{11}$, qui est non nul. Dans les formules de Cramer, il faut bien sûr remplacer la division par $\overline{11}$ par la multiplication par son inverse, qui est égal à $-\overline{10}$ ($110 = 3 \times 37 - 1$).*

SF1.12 Utiliser le théorème des restes chinois

Exemple

Résoudre l'équation $X^2 - \overline{6}X + \overline{5} = \overline{0}$ dans $\mathbb{Z}/91\mathbb{Z}$.

Cette équation peut aussi s'écrire : $(X - \overline{1})(X - \overline{5}) = \overline{0}$.
$\overline{1}$ *et* $\overline{5}$ *sont donc solutions, mais ce ne sont pas les seules puisque* 91 *n'est pas premier (*$91 = 7 \times 13$*).*
En notant $X = \overline{x}$ *avec* $x \in \mathbb{Z}$*, l'équation équivaut donc à :* 91 *divise* $(x-1)(x-5)$*. On a donc* $\{7 \mid (x-1) \text{ ou } 7 \mid (x-5)\}$ *et* $\{13 \mid (x-1) \text{ ou } 13 \mid (x-5)\}$.
Outre les solutions $x \equiv 1 \pmod{91}$ *et* $x \equiv 5 \pmod{91}$*, on obtient les deux systèmes :*

$$\begin{cases} x \equiv 1 \pmod 7 \\ x \equiv 5 \pmod{13} \end{cases} \quad \text{ou} \quad \begin{cases} x \equiv 5 \pmod 7 \\ x \equiv 1 \pmod{13} \end{cases} .$$

D'après le théorème chinois, il suffit de trouver une solution particulière dans $[\![0\,;90]\!]$ *de chacun de ces systèmes ; en essayant les congruences modulo* 13*, on trouve facilement qu'une solution particulière du 1er système est* $x_0 = 57$ *et pour le second système* $x_0 = 40$*, les autres solutions dans* $\mathbb{Z}$ *leur étant congrues modulo* 91.

En conclusion l'ensemble des solutions de l'équation proposée est :

$$\mathscr{S} = \{\overline{1}, \overline{5}, \overline{40}, \overline{57}\} .$$

SF1.13 Utiliser le théorème d'Euler ou de Fermat

Exemple

Quels sont les deux derniers chiffres de l'écriture de 3^{2016} en base 10 ?

Puisque 3 *est premier avec* 100*, le théorème d'Euler donne :*

$$3^{\varphi(100)} \equiv 1 \pmod{100} .$$

Or $100 = 2^2.5^2$ *donc* $\varphi(100) = 100\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{5}\right) = 40$*. On en déduit :*

$$3^{2016} = 3^{40.50+16} \equiv 3^{16} = 81^4 \equiv 19^4 \equiv 61^2 \equiv 21 \pmod{100} .$$

Les deux derniers chiffres de l'écriture décimale de 3^{2016} *sont donc* 21.

Exemple

Trouver le reste dans la division de 51^{24} par 72.

On a $\varphi(72) = 24$; mais 51 et 72 ne sont pas premiers entre eux, donc on ne peut pas appliquer ici le théorème d'Euler.
Cependant, d'après le théorème des restes chinois, pour connaître la classe de 51^{24} modulo 72, il suffit de connaître ses classes modulo 8 et modulo 9.

Or $\varphi(8) = 4$ et 51 est premier avec 8, donc le théorème d'Euler donne : $51^4 \equiv 1 \pmod 8$ donc $51^{24} \equiv 1 \pmod 8$.

D'autre part, puisque 51 est multiple de 3, on a $51^{24} \equiv 0 \pmod 9$.

D'après le théorème des restes chinois, $51^{24} \equiv 9 \pmod{72}$ (puisque $9 \equiv 1 \pmod 8$ et $9 \equiv 0 \pmod 9$) : le reste cherché est égal à 9.

■ Comment utiliser l'indicateur d'Euler ?

La formule qui donne la valeur de l'indicateur d'Euler peut être utilisée pour étudier des congruences grâce au théorème d'Euler, comme dans l'exemple précédent. Mais il ne faut pas oublier la définition initiale : $\varphi(n)$ est le nombre d'éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$, ou encore le nombre de générateurs d'un groupe cyclique d'ordre n. Ce sont ces deux notions qui sont les plus utiles dans les exercices théoriques.

SF1.14 Considérer les éléments inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$

Exemple

1) Soient a et p des entiers naturels non nuls tels que $a^{p-1} \equiv 1 \pmod p$.
On suppose que pour tout facteur premier q de $p-1$, $a^{\frac{p-1}{q}} \not\equiv 1 \pmod p$.
Montrer que p est premier (c'est le *test de primalité de Lucas-Lehmer*).

2) Réciproquement, montrer que si p est premier, il existe un entier a qui satisfait aux conditions précédentes.

1) *On considère l'ensemble $\mathcal{U}(\mathbb{Z}/p\mathbb{Z})$ des éléments inversibles de l'anneau $\mathbb{Z}/p\mathbb{Z}$. Son cardinal est $\varphi(p)$, où φ désigne comme d'habitude l'indicateur d'Euler.*
On sait que p est premier si et seulement si $\mathbb{Z}/p\mathbb{Z}$ est un corps, autrement dit si $\varphi(p) = p-1$.
Par hypothèse on a, dans $\mathbb{Z}/p\mathbb{Z}$, $\overline{a}^{p-1} = \overline{1}$; donc $\overline{a}$ est inversible et par suite son ordre d divise $p-1$. Mais $\overline{a}^{\frac{p-1}{q}} \neq \overline{1}$ donc d ne divise aucun des $\frac{p-1}{q}$ pour tout facteur premier q de $p-1$. Il en résulte que $d = p-1$ donc $\varphi(p) \geqslant p-1$ (les $p-1$ puissances de $\overline{a}$ étant toutes inversibles). Finalement, $\varphi(p) = p-1$ et p est premier.

2) *Réciproquement, si p est premier, tout générateur a du groupe $\mathcal{U}(\mathbb{Z}/p\mathbb{Z})$ des éléments inversibles de l'anneau $\mathbb{Z}/p\mathbb{Z}$ vérifie les conditions voulues.*

SF1.15 Considérer les générateurs d'un groupe cyclique

Exemple

Soit p un nombre premier. Démontrer que le groupe $\mathcal{U}(\mathbb{Z}/p\mathbb{Z})$ des éléments inversibles de l'anneau $\mathbb{Z}/p\mathbb{Z}$ est cyclique.

Pour simplifier, notons $G = \mathcal{U}(\mathbb{Z}/p\mathbb{Z})$. G est donc un groupe d'ordre $p-1$. Soit $a \in G$, d'ordre d, avec $d \mid (p-1)$. Le sous-groupe engendré par a est $\mathrm{gr}(a) = \{\overline{1}, a, \ldots, a^{d-1}\}$, et il possède $\varphi(d)$ générateurs, qui sont donc des éléments de G d'ordre d.

Puisque $\mathbb{Z}/p\mathbb{Z}$ est un corps, le polynôme $X^d - \overline{1} \in \mathbb{Z}/p\mathbb{Z}[X]$ admet au plus d racines, et comme tout élément de $\mathrm{gr}(a)$ en est une racine, les racines de $X^d - \overline{1}$ dans $\mathbb{Z}/p\mathbb{Z}$ sont exactement les éléments de $\mathrm{gr}(a)$.
Ainsi, si x est un élément de G d'ordre d, il est racine dans $\mathbb{Z}/pZ$ du polynôme $X^d - \overline{1}$ donc x appartient à $\mathrm{gr}(a)$.
Les éléments de G d'ordre d, s'il en existe, sont donc ceux de $\mathrm{gr}(a)$, donc il y en a $\varphi(d)$.

Pour tout diviseur d de $p-1$, notons n_d le nombre d'éléments de G d'ordre d. On vient de voir que $n_d = 0$ ou $n_d = \varphi(d)$. Or

$$p - 1 = \mathrm{Card}(G) = \sum_{d \mid (p-1)} n_d$$

puisque tout élément de G a un ordre qui divise $p-1$, et l'on sait aussi d'après le cours que :

$$p - 1 = \sum_{d \mid (p-1)} \varphi(d)\,.$$

Il en résulte que $n_d = \varphi(d)$ pour tout diviseur d de $p-1$. n_d n'est donc jamais nul et, en particulier, il existe dans G un élément d'ordre $p-1$, ce qui signifie que G est cyclique.

■ Comment résoudre un problème mettant en jeu des polynômes ?

Contrairement à ce que l'on pourrait croire, les exercices sur les polynômes sont souvent loin d'être faciles. Bien sûr, vous connaissez (ou croyez connaître) les polynômes depuis longtemps puisqu'il s'agit des premières fonctions que l'on étudie dans le secondaire.
Mais justement, un polynôme ne se réduit pas à une simple fonction polynôme, loin de là ! Un problème mettant en jeu des polynômes peut être abordé sous l'aspect algébrique (structure d'anneau principal, divisibilité, factorisation, ...), l'aspect linéaire (structure d'espace vectoriel, utilisation de bases, ...), l'aspect bilinéaire (produit scalaire, polynômes orthogonaux,...) ou l'aspect fonctionnel (continuité, racines, variations, ...).
Nous essayons donc de donner ci-dessous une première liste de méthodes parmi les plus courantes, sans vouloir être exhaustif ; cette liste sera complétée dans les chapitres suivants. Enfin, il faut savoir que les problèmes les plus délicats devront utiliser conjointement plusieurs de ces méthodes.

SF1.16 Utiliser le degré

La détermination du degré d'un polynôme est en général chose facile et peut parfois apporter un renseignement important.

✓ Considérer les degrés des polynômes qui interviennent dans un problème doit être une démarche presque systématique.

Exemple

Déterminer tous les polynômes $P \in \mathbb{K}[X]$ tels que :

$$X(X-1)P' + P^2 - (2X+1)P + 2X = 0\,.$$

Le polynôme nul n'est pas solution.
Soit P un polynôme solution de degré $n \in \mathbb{N}$. Si $n \geqslant 2$, alors $\deg(P^2) = 2n$ et $\deg(X(X-1)P' - (2X+1)P + 2X) \leqslant n+1 < 2n$, c'est impossible.
Donc nécessairement $\deg P \leqslant 1$, et en remplaçant P par $aX+b$ dans l'équation, on trouve comme solutions :

$$P \in \{1,\, X,\, X+1\}\,.$$

SF1.17 Utiliser les coefficients

Si nous avons placé cette méthode au début de ce paragraphe, c'est parce que la plupart d'entre vous ont un très mauvais réflexe en abordant un exercice mettant en jeu des polynômes : commencer par écrire chaque polynôme sous la forme $\sum a_k X^k \ldots$
Or il s'agit non pas de la première piste à envisager, mais souvent de la dernière, car elle peut conduire à des calculs inextricables. Bien sûr, les coefficients sont le point de départ de la construction de l'algèbre $\mathbb{K}[X]$, mais le reste du cours contient des techniques bien plus performantes.
Cette méthode est donc à réserver, en général, aux cas où les polynômes considérés sont de degrés « petits ».

Exemple

Trouver tous les polynômes $P \in \mathbb{R}[X]$ tels que $4P = (X-1)P' + P''$.

- *Le seul polynôme constant solution est le polynôme nul.*
  Si P n'est pas constant et si $n = \deg P$, il est facile de vérifier, en comparant les coefficients dominants des deux membres de l'égalité, que l'on a nécessairement $n = 4$.
- *On peut donc chercher P sous la forme :*

  $$P = aX^4 + bX^3 + cX^2 + dX + e\,.$$

  Après de passionnants calculs, on trouve que P est solution si et seulement si ses coefficients sont solutions du système :

  $$\begin{cases} 4a + b = 0 \\ -12a + 3b + 2c = 0 \\ -6b + 2c + 3d = 0 \\ -2c + d + 4e = 0 \end{cases},$$

et la résolution tout aussi passionnante de ce système donne :

$$a = \frac{1}{10}e \ , \ b = -\frac{2}{5}e \ , \ c = \frac{6}{5}e \ \text{ et } \ d = -\frac{8}{5}e \, .$$

Les polynômes P solutions sont donc ceux de la forme :

$$P = \lambda(X^4 - 4X^3 + 12X^2 - 16X + 10) \ , \ \lambda \in \mathbb{R} \, .$$

✓ Nous vous avons montré ci-dessus ce qu'il ne faut *pas* faire (c'est d'ailleurs pour cette raison que nous n'avons pas détaillé les calculs). On pouvait en effet résoudre le problème bien plus facilement en exprimant P non pas dans la base canonique, mais dans la base formée des polynômes $(X-1)^k$ avec $0 \leqslant k \leqslant 4$. Cette méthode permet d'ailleurs de résoudre l'exercice en remplaçant « 4 » par n'importe quelle autre valeur.

SF1.18 Utiliser la divisibilité dans $\mathbb{K}[X]$

Outre le célèbre théorème de Gauss, on utilise aussi fréquemment deux autres résultats importants.

- Si P est un polynôme irréductible qui divise un produit AB, alors P divise A ou P divise B .
- Si A et B divisent P et sont premiers entre eux ; leur produit AB divise P.

Exemple

Déterminer tous les polynômes P, Q, $R \in \mathbb{K}[X]$ tels que :

$$(X^3+1)P^2 - XQ^2 = X^3R^2 - 2X^2QR \, .$$

Soient P, Q, $R \in \mathbb{K}[X]$ vérifiant la relation proposée.
Alors X divise $(X^3+1)P^2$; étant premier avec X^3+1, X divise P^2 (théorème de Gauss), et puisqu'il est irréductible, X divise P.
Mais alors X^2 divise P^2 donc divise XQ^2 d'où X divise Q^2 puis X divise Q.
Posons donc $Q = XQ_1$ et $P = XP_1$. La relation donnée équivaut alors à :

$$(X^3+1)P_1^2 = X(R^2 - 2Q_1R + Q_1^2) = X(R-Q_1)^2 \, .$$

Si ce polynôme n'est pas nul, la valuation de X dans $X(R-Q_1)^2$ est impaire alors que dans $(X^3+1)P^2$, elle est paire : c'est impossible.

Il en résulte que $P_1 = 0$ et $R = Q_1$, soit $P = 0$ et $Q = XR$.

Réciproquement, il est facile de vérifier que tous les triplets de polynômes $(P,Q,R) = (0,XR,R)$ avec $R \in \mathbb{K}[X]$ quelconque, conviennent.

SF1.19 Utiliser la relation ou le théorème de Bezout

Exemple

Soient A, $B \in \mathbb{R}[X]$ tels que : $1 - A^2 = (1-X^2)B^2$, et soit $n = \deg(A)$ $(n \geqslant 1)$.

Établir que $A' = \pm nB$, et en déduire une équation différentielle linéaire du second ordre vérifiée par A.

- *La relation* $A^2 + (1 - X^2)B^2 = 1$ *peut aussi s'écrire* $UA + VB = 1$ *avec* $U = A$ *et* $V = (1 - X^2)B$, *donc elle implique, par le théorème de Bezout, que* $A \wedge B = 1$.
- *En dérivant cette relation on obtient :*
$$-AA' = (1 - X^2)BB' - XB^2 = \big((1 - X^2)B' - XB\big)B\,.$$
Ainsi B *divise le produit* AA' *et puisqu'il est premier avec* A, *il divise* A' *d'après le théorème de Gauss.*
La relation $1 - A^2 = (1 - X^2)B^2$ *implique* $\deg B = \deg A - 1 = \deg A'$, *donc il existe* $\lambda \in \mathbb{R}^*$ *tel que* $A' = \lambda B$ *(*A *n'est pas constant par hypothèse). Si le coefficient dominant de* A *est* a_n, *celui de* A' *est égal à* na_n *; d'autre part, d'après la relation de l'énoncé, le coefficient dominant de* B^2 *est égal à* a_n^2, *donc celui de* B *est égal à* $\pm a_n$, *d'où* $\lambda = \pm n$.
On a donc montré que $A' = \pm nB$.
- *On a donc la relation :* $A^2 + \dfrac{1}{n^2}(1 - X^2)A'^2 = 1$, *puis en la dérivant on obtient :*
$$AA' + \frac{1}{n^2}\left((1 - X^2)A'A'' - XA'^2\right) = 0\,.$$
Puisque $A' \neq 0$ *et que l'anneau* $\mathbb{R}[X]$ *est intègre on en tire :*
$$(1 - X^2)A'' - XA' + n^2A = 0\,.$$

✓ *La résolution de cette équation différentielle, à l'aide du changement de variable* $x = \cos t$, *permet de montrer que* $A = \pm T_n$ *où* T_n *est le* n^e *polynôme de Tchebychev de 1re espèce, défini par* $\cos nt = T_n(\cos t)$.

SF1.20 Utiliser l'algorithme d'Euclide

L'algorithme d'Euclide, vu en 1re année, permet de déterminer le pgcd de deux polynômes A et B, mais aussi, en combinant les relations obtenues lors des divisions successives, de déterminer deux polynômes U et V tels que $AU + BV = A \wedge B$.
Nous ne donnerons pas ici d'exemple calculatoire, nous vous conseillons de revoir le cours de 1re année à ce sujet.

Exemple

Soient a et b deux nombres entiers non nuls, et $d = a \wedge b$.

Montrer que le pgcd dans $\mathbb{R}[X]$ des polynômes $X^a - 1$ et $X^b - 1$ est égal à $X^d - 1$.

- *La division euclidienne de* a *par* b *s'écrit :*
$$a = bq + r \text{ avec } 0 \leqslant r < b\,.$$
On a alors : $X^a - 1 = (X^{bq} - 1)X^r + X^r - 1$.
Puisque $X^{bq} - 1 = (X^b)^q - 1$ *est divisible par* $X^b - 1$ *et que* $\deg(X^r - 1) < \deg(X^b - 1)$, *la relation précédente implique que* $X^r - 1$ *est le reste dans la division euclidienne de* $X^a - 1$ *par* $X^b - 1$.
- *On a alors* $a \wedge b = b \wedge r$ *et* $X^a - 1 \wedge X^b - 1 = X^b - 1 \wedge X^r - 1$.
- *On effectue alors la suite des divisions euclidiennes comme dans l'algorithme d'Euclide. Le pgcd de* a *et* b *est égal au dernier reste* r_n *non nul :*
$$a \wedge b = b \wedge r = r \wedge r_1 = \ldots = r_{n-1} \wedge r_n = r_n \wedge r_{n+1} \quad \text{et} \quad r_{n+1} = 0\,.$$

On aura donc de la même façon :

$$X^a-1\wedge X^b-1=X^b-1\wedge X^r-1=\ldots=X^{r_n}-1\wedge\underbrace{X^{r_{n+1}}-1}_{=0},$$

donc : $X^a-1\wedge X^b-1=X^{r_n}-1=X^d-1$.

SF1.21 Utiliser l'arithmétique dans $\mathbb{Z}$

Il est naturel d'utiliser l'arithmétique dans $\mathbb{Z}$ pour obtenir des propriétés sur un polynôme appartenant à $\mathbb{Z}[X]$. Et cela permet aussi d'obtenir des propriétés sur les polynômes à coefficients rationnels, car si P appartient à $\mathbb{Q}[X]$, en le multipliant par le ppcm des dénominateurs de ses coefficients, on obtient un polynôme à coefficients entiers.
L'exemple ci-dessous est bien utile pour essayer de trouver des racines « évidentes » à un polynôme.

Exemple

Soit $P=a_nX^n+\cdots+a_1X+a_0\in\mathbb{Z}[X]$.

Montrer que si un nombre rationnel $r=\dfrac{p}{q}$ écrit sous forme irréductible est racine de P, alors $p\big|a_0$ et $q\big|a_n$.

Application : trouver la racine rationnelle du polynôme $P=12X^3-8X^2+7X-1$.

- *Supposons* $r=\dfrac{p}{q}$ *racine de* P, *avec* $p\wedge q=1$.
  L'égalité $P(r)=0$ *s'écrit aussi :*
  $$a_np^n+a_{n-1}p^{n-1}q+\cdots+a_1pq^{n-1}+a_0q^n=0\,.$$
  On en déduit que p *divise* a_0q^n *; puisque* $p\wedge q=1$, *on a aussi* $p\wedge q^n=1$, *et le théorème de Gauss implique que* p *divise* a_0.
  De même, a_np^n *est multiple de* q, *donc* q *divise* a_n.
- *Si* $P=12X^3-8X^2+7X-1$ *admet une racine rationnelle* $r=\dfrac{p}{q}$, *on a nécessairement* $p\big|1$ *et* $q\big|12$ *donc :*
  $$r\in\left\{\pm\frac{1}{12},\ \pm\frac{1}{6},\ \pm\frac{1}{4},\ \pm\frac{1}{3},\ \pm\frac{1}{2},\ \pm1\right\}.$$
  Mais pour $x\leqslant0$, $P(x)<0$, *donc les solutions négatives sont à éliminer. Il ne reste plus que* 6 *valeurs à essayer, et finalement on trouve* $r=\dfrac{1}{6}\cdot$
  On en déduit $P=(6X-1)(2X^2-X+1)$, *donc* $\dfrac{1}{6}$ *est la seule racine réelle de* P.

SF1.22 Utiliser la décomposition en facteurs irréductibles

Exemple

Soit P un polynôme à coefficients réels tel que pour tout x réel on ait $P(x)\geqslant0$.

1) Que peut-on dire sur le degré de P ?

2) Montrer que les racines réelles de ce polynôme sont d'ordre de multiplicité paire.

3) Montrer que P peut s'écrire sous la forme $P=A^2+B^2$ où A et B sont deux polynômes à coefficients réels.

1) *Une fonction polynôme de degré impair change nécessairement de signe (considérer ses limites en $\pm\infty$). Donc si P vérifie la condition de l'énoncé,* $\deg P$ *est pair.*

2) *Si x_0 est une racine réelle de P d'ordre de multiplicité m, on a, d'après la formule de Taylor :*

$$P(x) \underset{x\to x_0}{\sim} \frac{P^{(m)}(x_0)}{m!}(x-x_0)^m\,,$$

donc si m est impair, il existe un voisinage de x_0 où P change de signe, d'où le résultat.

3) *Puisque les (éventuelles) racines complexes non réelles de P sont deux à deux conjuguées et ont même ordre de multiplicité, la décomposition de P en facteurs irréductibles dans $\mathbb{C}[X]$ s'écrit :*

$$\begin{aligned} P &= \prod_{i=1}^{r}(X-a_i)^{2\alpha_i}\prod_{j=1}^{s}(X-z_j)^{\beta_j}(X-\overline{z_j})^{\beta_j} \\ &= \left(\prod_{i=1}^{r}(X-a_i)^{\alpha_i}\prod_{j=1}^{s}(X-z_j)^{\beta_j}\right)\left(\prod_{i=1}^{r}(X-a_i)^{\alpha_i}\prod_{j=1}^{s}(X-\overline{z_j})^{\beta_j}\right), \end{aligned}$$

(où $a_i\in\mathbb{R}$, $z_j\in\mathbb{C}\setminus\mathbb{R}$, et certains exposants α_i ou β_j peuvent être nuls). P peut donc s'écrire sous la forme :

$$P = Q\cdot\overline{Q} = (A+\mathrm{i}B)(A-\mathrm{i}B) = A^2+B^2$$

où $Q\in\mathbb{C}[X]$ et A, $B\in\mathbb{R}[X]$.

SF1.23 Utiliser les racines

On utilise régulièrement le fait que a est racine d'un polynôme P si et seulement si $(X-a)$ divise P, ou plus généralement : a est racine de P d'ordre supérieur ou égal à m si et seulement si $(X-a)^m$ divise P.
Une autre propriété fréquemment utilisée est que si un polynôme possède une infinité de racines (ou au moins strictement plus de racines que son degré), c'est le polynôme nul.

Exemple

Déterminer tous les polynômes $P\in\mathbb{K}[X]$ tels que : $(X+3)P(X) = XP(X+1)$.

On voit facilement que $P(0)=0$ puis $P(-1)=0$ puis $P(-2)=0$.
Ainsi $P = X(X+1)(X+2)Q$. En remplaçant, on trouve $Q(X)=Q(X+1)$, d'où $Q=$ cste (car le polynôme $Q-Q(0)$ a tous les entiers pour racines).
Donc $P(X)=\lambda X(X+1)(X+2)$ avec $\lambda\in\mathbb{K}$. Réciproquement, on vérifie aisément que tous ces polynômes conviennent.

Dans $\mathbb{C}[X]$, l'utilisation des racines simplifie beaucoup l'étude de la divisibilité : en effet, d'après la décomposition d'un polynôme en produit de facteurs irréductibles de $\mathbb{C}[X]$, un polynôme P divise un polynôme Q dans $\mathbb{C}[X]$ si et seulement si toutes les racines de P sont aussi racines de Q avec un ordre de multiplicité supérieur.

Exemple

Déterminer les $n \in \mathbb{N}^*$ pour lesquels le polynôme $A = (X^2+X+1)^2$ divise le polynôme $B_n = (X+1)^n - X^n - 1$ dans $\mathbb{R}[X]$.

Les deux polynômes étant à coefficients réels, A divise B_n dans $\mathbb{R}[X]$ si et seulement si il le divise dans $\mathbb{C}[X]$.
Les racines de A dans $\mathbb{C}$ étant j et j^2, et étant d'ordre 2, A divise B_n si et seulement si j et j^2 sont racines de B_n d'ordre supérieur ou égal à 2 ; B_n étant à coefficients réels, il suffit de faire la vérification pour j.
Cela équivaut donc à $B_n(\mathrm{j}) = B_n'(\mathrm{j}) = 0$, c'est-à-dire à :

$$(\mathrm{j}+1)^n - \mathrm{j}^n = 1 \quad \text{et} \quad (\mathrm{j}+1)^{n-1} = \mathrm{j}^{n-1}\,.$$

Puisque $\mathrm{j}+1 = -\mathrm{j}^2$, la deuxième équation équivaut à : $(-\mathrm{j})^{n-1} = 1$, soit à $\mathrm{e}^{-\frac{\mathrm{i}(n-1)\pi}{3}} = 1$ ou encore à $n \equiv 1 \pmod 6$. Dans ce cas, puisque $\mathrm{j}^3 = 1$,

$$(\mathrm{j}+1)^n - \mathrm{j}^n = (-\mathrm{j}^2)^n - \mathrm{j}^n = -j^2 - j = 1\,,$$

donc la 1re équation est automatiquement vérifiée.
En conclusion, A divise B_n si et seulement si $n \equiv 1 \pmod 6$.

Il est également important de se rappeler que a est racine d'ordre m de P si et seulement si c'est une racine d'ordre $m-1$ de P' **et** si $P(a) = 0$ (ne pas oublier cette dernière condition).

Exemple

Déterminer tous les polynômes $P \in \mathbb{R}[X]$ tels que $(X+1)^4$ divise $P-1$ et $(X-1)^4$ divise $P+1$.

✓ *Il s'agit ici d'un problème analogue au problème des restes chinois dans $\mathbb{Z}$, puisqu'il s'agit de résoudre le système :*

$$P \equiv 1 \pmod{(X+1)^4} \quad \text{et} \quad P \equiv -1 \pmod{(X-1)^4}\,.$$

- *Comme dans le cas de l'arithmétique dans $\mathbb{Z}$, cherchons d'abord une solution particulière P_0.*
  Si P_0 convient, -1 est racine d'ordre 4 au moins de $P_0 - 1$, donc racine d'ordre 3 au moins de P_0'.
  De même, 1 est racine d'ordre 3 au moins de P_0', et P_0' est donc divisible par $(X-1)^3(X+1)^3$.
  De plus, -1 doit être racine de $P_0 - 1$ d'où $P_0(-1) = 1$ et 1 est racine de $P_0 + 1$ soit $P_0(1) = -1$. Puisqu'il y a deux conditions, on va chercher un polynôme P_0 de degré $\leqslant 7$.
  P_0' étant de degré $\leqslant 6$, il existe un réel a tel que :

$$P' = a(X-1)^3(X+1)^3 = a(X^6 - 3X^4 + 3X^2 - 1)\,,$$

  donc il existe b tel que $P = a\left(\frac{X^7}{7} - \frac{3}{5}X^5 + X^3 - X\right) + b$.
  Les conditions $P_0(-1) = 1$ et $P_0(1) = -1$ fournissent les équations :

$$\frac{16}{35}a + b = 1 \quad \text{et} \quad -\frac{16}{35}a + b = -1\,,$$

d'où $b = 0$ *et* $a = \dfrac{35}{16}$.

Finalement : $P_0 = \dfrac{1}{16}\left(5X^7 - 21X^5 + 35X^3 - 35X\right)$ *(ce polynôme convient effectivement car toutes les conditions écrites précédemment sont en fait nécessaires et suffisantes).*

- *Soit maintenant* P *un polynôme quelconque solution. Alors :*

$$\begin{aligned}(X+1)^4 \big| (P-1) &\iff (X+1)^4 \big| (P-1) \quad \text{et} \quad (X+1)^4 \big| (P_0 - 1) \\ &\iff (X+1)^4 \big| (P - P_0) \quad \text{et} \quad (X+1)^4 \big| (P_0 - 1) \\ &\iff (X+1)^4 \big| (P - P_0)\,,\end{aligned}$$

et de même $(X-1)^4 \big| (P+1) \iff (X-1)^4 \big| (P - P_0)$.
$(X-1)^4$ *et* $X+1)^4$ *étant premiers entre eux,* P *est solution si et seulement si* $(X-1)^4(X+1)^4$ *divise* $P - P_0$.

Finalement les solutions au problème sont les polynômes de la forme :

$$P = \frac{1}{16}\left(5X^7 - 21X^5 + 35X^3 - 35X\right) + (X-1)^4(X+1)^4 Q\,,$$

où Q *est un polynôme quelconque de* $\mathbb{R}[X]$.

SF1.24 Utiliser les relations coefficients-racines

Exemple

Soit $P = X^4 - 5X^3 + 9X^2 - 15X + 18$.

Trouver les racines de P sachant que le produit de deux d'entre elles vaut 6.

En notant x_i *les racines* $(1 \leqslant i \leqslant 4)$, $s_1 = x_1 + x_2$, $s_2 = x_3 + x_4$, $p_1 = x_1 x_2$ *et* $p_2 = x_3 x_4$, *les relations coefficients-racines donnent :*

$$\begin{array}{lcll} \sigma_1 & = & s_1 + s_2 & = & 5 \\ \sigma_2 & = & p_1 + p_2 + s_1 s_2 & = & 9 \\ \sigma_3 & = & p_1 s_2 + p_2 s_1 & = & 15 \\ \sigma_4 & = & p_1 p_2 & = & 18 \end{array}$$

et on trouve, puisque $p_1 = 6$, $p_2 = 3$ *puis* $s_1 = 5$ *et* $s_2 = 0$.

Finalement, les racines cherchées sont 2, 3, $\mathrm{i}\sqrt{3}$, $-\mathrm{i}\sqrt{3}$.

✓ *Pour obtenir ce dernier résultat, nous avons utilisé le fait que si l'on connaît la somme* s *et le produit* p *de deux nombres, ceux-ci sont solutions de l'équation* $X^2 - sX + p = 0$.

Dans l'exemple précédent, l'utilisation des relations coefficients-racines était pratiquement imposée par l'énoncé. Mais il y a des cas où, pour les utiliser, il faut penser à introduire un polynôme adéquat.

Exemple

Soient a, b, c trois nombres complexes de modules distincts, tels que :

$$\forall\, k \in \{1,2,3\},\ a^k + b^k + c^k \in \mathbb{R}\,.$$

Montrer que a, b et c sont réels.

a, b, c sont les racines de $P = X^3 - \sigma_1 X^2 + \sigma_2 X - \sigma_3$, où $\sigma_1 = a+b+c$, $\sigma_2 = ab+bc+ca$ et $\sigma_3 = abc$.
On a, en notant $\alpha_k = a^k + b^k + c^k$ pour $k \in \{1,2,3\}$:

$$\sigma_1 = \alpha_1 \ , \ 2\sigma_2 = \alpha_1^2 - \alpha_2 \quad et \quad 3\sigma_3 = \alpha_3 - \sigma_1\alpha_2 + \sigma_2\alpha_1 \, ,$$

cette dernière égalité étant obtenue en additionnant les trois égalités $P(a) = P(b) = P(c) = 0$.
L'hypothèse de l'énoncé implique que les σ_k sont réels, donc $P \in \mathbb{R}[X]$.
P étant de degré impair admet au moins une racine réelle, par exemple a.
Si b et c étaient complexes non réels, on aurait $b = \overline{c}$ d'où $|b| = |c|$, ce qui est exclu.
En conclusion, les trois racines de P, a, b et c, sont réelles.

✓ Il faut penser à utiliser les relations coefficients-racines pour un polynôme de degré n lorsque l'énoncé fait apparaître des fonctions symétriques de n variables.

SF1.25 Utiliser la formule de Taylor

La formule de Taylor est bien sûr intéressante dès qu'apparaît une racine et son ordre de multiplicité, mais il est important de retenir aussi qu'elle permet d'obtenir les coefficients d'un polynôme P dans la base de $\mathbb{K}[X]$ formée des $(X-a)^k$: ce sont les scalaires $\dfrac{P^{(k)}(a)}{k!}$.
En particulier, les coefficients d'un polynôme dans la base canonique sont les $\dfrac{P^{(k)}(0)}{k!}$.

Exemple

Déterminer les polynômes P non constants tels que P' divise P.

Soit $n = \deg P$. Si P' divise P, le quotient est nécessairement de degré 1 donc il existe un réel a tel que :

$$P = \frac{1}{n}(X-a)P'$$

(le coefficient $\dfrac{1}{n}$ a été obtenu en comparant les termes dominants).
On applique alors la formule de Taylor :

$$P = \sum_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k \ , \ d'où \ P' = \sum_{k=1}^{n} \frac{kP^{(k)}(a)}{k!}(X-a)^{k-1} \, .$$

La relation $(X-a)P' = nP$ équivaut alors, par égalité des coordonnées des polynômes dans la base formée des $(X-a)^k$, à :

$$\forall\, k \in [\![0\,;n]\!], \ k\frac{P^{(k)}(a)}{k!} = n\frac{P^{(k)}(a)}{k!} \, ,$$

donc à :

$$\forall\, k \in [\![0\,;n-1]\!], \ P^{(k)}(a) = 0 \, .$$

a est donc racine de P d'ordre n, c'est-à-dire que P est de la forme $\lambda(X-a)^n$ avec $\lambda \in \mathbb{R}$.
Il est facile de vérifier enfin que tous ces polynômes conviennent.

SF1.26 Utiliser la fraction rationnelle P'/P

Si P est un polynôme de $\mathbb{K}[X]$ scindé, il s'écrit sous la forme :

$$P = \lambda \prod_{i=1}^{r} (X - a_i)^{\alpha_i} .$$

Alors par un simple calcul du polynôme dérivé on a :

$$\frac{P'}{P} = \sum_{i=1}^{r} \frac{\alpha_i}{X - a_i}$$

(il s'agit de la décomposition en éléments simples de la fraction rationnelle P'/P).

Exemple

On désigne par $a_1, \ldots, a_5$ les racines complexes du polynôme $P = X^5 + 2X^4 - X - 1$.

Calculer la somme : $\displaystyle\sum_{k=1}^{5} \frac{a_k + 2}{a_k - 1}$.

D'après la remarque précédente on a $\dfrac{P'}{P} = \displaystyle\sum_{k=1}^{5} \frac{1}{X - a_k}$.

On écrit alors $a_k + 2 = (a_k - 1) + 3$ *d'où :*

$$\sum_{k=1}^{5} \frac{a_k + 2}{a_k - 1} = 5 - 3 \sum_{k=1}^{5} \frac{1}{1 - a_k} = 5 - 3\frac{P'(1)}{P(1)} = 5 - 36 = -31.$$

Exemple

Soit P un polynôme non constant de $\mathbb{C}[X]$.
Démontrer que l'image dans le plan complexe des racines de P' appartient à l'enveloppe convexe des images des racines de P (c'est le *théorème de Lucas*).

✓ *Rappelons que l'enveloppe convexe d'une famille de points $A_1, \ldots, A_n$ est l'ensemble des barycentres des A_i à coefficients positifs (voir plus de détail en* **[S9.2]**).

$P \in \mathbb{C}[X]$ *est scindé :* $P = \lambda \displaystyle\prod_{i=1}^{r} (X - z_i)^{\alpha_i}$. *On a alors* $\dfrac{P'}{P} = \displaystyle\sum_{i=1}^{r} \frac{\alpha_i}{X - z_i}$.

Soit Z une racine de P' dans $\mathbb{C}$. On envisage deux cas :

– *si Z est aussi racine de P, le résultat est immédiat.*

– *sinon, on peut écrire :*

$$0 = \frac{P'(Z)}{P(Z)} = \sum_{i=1}^{r} \frac{\alpha_i}{Z - z_i} = \sum_{i=1}^{r} \frac{\alpha_i(\overline{Z} - \overline{z_i})}{|Z - z_i|^2},$$

d'où en prenant le conjugué :

$$\sum_{i=1}^{r} \frac{\alpha_i(Z - z_i)}{|Z - z_i|^2} = 0 .$$

Cela signifie que Z est barycentre des z_i affectés des coefficients $\dfrac{\alpha_i}{|Z - z_i|^2}$, *qui sont des réels positifs, d'où le résultat.*

SF1.27 Considérer la fonction polynôme associée

Dans le cas de polynômes à coefficients réels, on peut souvent utiliser des outils de l'analyse : étude des variations, théorème des valeurs intermédiaires, théorème de Rolle etc...

Exemple

Soit $n \in \mathbb{N}^*$ et $P = X^n - \sum_{i=0}^{n-1} a_i X^i \in \mathbb{R}[X]$, les a_i étant positifs non tous nuls.

Montrer que P n'a qu'une et une seule racine réelle strictement positive, et que cette racine est simple (on pourra raisonner par récurrence).

- *Le résultat est immédiat si $n = 1$.*
- *Supposons le résultat démontré à l'ordre n, et soit $P = X^{n+1} - \sum_{i=0}^{n} a_i X^i$, les a_i étant positifs non tous nuls.*
  - *Si tous les a_i pour $i \geqslant 1$ sont nuls, alors $P = X^{n+1} - a_0$ avec $a_0 > 0$, et le résultat est acquis.*
  - *Sinon, on peut appliquer l'hypothèse de récurrence au polynôme $\frac{1}{n+1}P' = X^n - \sum_{i=0}^{n-1} \frac{i+1}{n+1} a_{i+1} X^i$; ainsi, P' possède une et une seule racine strictement positive, a, qui est simple ; donc P' change de signe en a ; de plus, $\lim_{x \to +\infty} P'(x) = +\infty$, donc P' est < 0 sur $]0\,;a[$ et > 0 sur $]a\,;+\infty[$.*

    Puisque $P(0) = -a_0 \leqslant 0$, le tableau de variations de P montre alors que P possède une et une seule racine strictement positive b, et $b > a$. b n'étant pas racine de P', c'est une racine simple de P, ce qui démontre le résultat.

Exemple

Soit $P = \sum_{k=0}^{n} a_k X^k \in \mathbb{R}[X]$, de degré $n \geqslant 1$, scindé dans $\mathbb{R}[X]$.

1) Soit $\lambda \in \mathbb{R}_+^*$. Montrer que le polynôme $\lambda P + XP'$ est scindé sur $\mathbb{R}$.

2) Soit $Q \in \mathbb{R}[X]$ un polynôme scindé unitaire n'ayant que des racines strictement négatives.

Montrer que le polynôme $R = \sum_{k=0}^{n} a_k Q(k) X^k$ est scindé sur $\mathbb{R}$.

1) *Notons $x_1 < x_2 < \ldots < x_p$ les racines non nulles de P, $\alpha_1, \ldots, \alpha_p$ leurs ordres de multiplicité et β l'ordre de multiplicité de 0 comme racine de P ($\beta = 0$ si 0 n'est pas racine). On a donc $n = \beta + \sum_{i=1}^{p} \alpha_i$.*

Le polynôme $Q = \lambda P + XP'$ admet alors $x_1, \ldots, x_p$ comme racines d'ordres respectifs $\alpha_i - 1$, et 0 comme racine d'ordre β, ce qui fait $n - p$ racines.

Considérons maintenant la fraction rationnelle :

$$S = \frac{Q}{XP} = \frac{\lambda}{X} + \frac{P'}{P} = \sum_{i=1}^{p} \frac{\alpha_i}{X - x_i} + \frac{\lambda + \beta}{X},$$

où $\lambda + \beta$ est strictement positif compte tenu des hypothèses.
En examinant alors les limites à droite et à gauche aux bornes et d'après le théorème des valeurs intermédiaires, on voit que S admet (au moins) une racine dans chacun des intervalles ouverts délimités par $x_1, \ldots, x_p$ et 0, ce qui fait p intervalles donc p racines. Il en est donc de même de Q, ce qui fait bien n racines au total, et Q est scindé.

2) *On procède par récurrence sur le degré p de Q.*

- *Dans le cas $p = 1$, le polynôme Q s'écrit $Q = X + \lambda$ avec $\lambda > 0$. Le polynôme R est alors égal à $XP' + \lambda P$, et on applique alors le résultat précédent.*
- *Supposons le résultat démontré à l'ordre $p \geqslant 1$, et soit Q de degré $p + 1$ vérifiant les hypothèses de l'énoncé. On peut alors écrire $Q = (X + \lambda)Q_1$, avec $\lambda > 0$ et Q_1 de degré p, scindé, unitaire et à racines strictement négatives. On a donc :*

$$R = \sum_{k=0}^{n} a_k Q_1(k) k X^k + \lambda \sum_{k=0}^{n} a_k Q_1(k) X^k = X P_1' + \lambda P_1 ,$$

où $P_1 = \sum_{k=0}^{n} a_k Q_1(k) X^k$. D'après l'hypothèse de récurrence, P_1 est scindé dans $\mathbb{R}[X]$. En lui appliquant le résultat de la 1re question, on obtient R scindé, ce qui démontre le résultat.

✓ Notez dans cet exemple, comme dans d'autres précédents, l'importance du raisonnement par récurrence sur le degré des polynômes concernés. La difficulté est souvent d'y penser !

À vous de jouer !

Exercice 1.1 **SF1.1**

1) Soit $(G,*)$ un groupe, et H une partie stable finie non vide de G.
Montrer que H est un sous-groupe de G.

2) Quelles sont les parties finies non vides de $\mathbb{C}^*$ stables pour la multiplication ?

Exercice 1.2 **SF1.5**

Soit G un groupe d'ordre $2p$, avec p premier.
Montrer que G contient un élément d'ordre p.

Exercice 1.3 **SF1.3** **SF1.4**

Soit G un groupe abélien noté multiplicativement. On suppose qu'il existe $n \in \mathbb{N}^*$ tel que $x^n = e_G$ pour tout $x \in G$.

1) On suppose que $n = ab$ avec $a \wedge b = 1$. On pose $G_a = \{x^a \mid x \in G\}$ et $G_b = \left\{x^b \;\middle|\; x \in G\right\}$.

a) Montrer que G_a et G_b sont des sous-groupes de G.

b) Montrer que pour tout $x \in G$ il existe un unique couple $(y,z) \in G_a \times G_b$ tel que $x = yz$.

2) Soit k un entier premier avec n.
Montrer que l'application $f_k : x \mapsto x^k$ est un automorphisme de G et préciser son application réciproque.

Indication pour les deux questions : penser à utiliser l'identité de Bezout.

Exercice 1.4 **SF1.9**

Un anneau commutatif A est dit local si l'ensemble $N(A)$ des éléments non inversibles de A est un idéal.

1) Montrer que l'anneau $\mathbb{Z}/n\mathbb{Z}$ $(n \geqslant 2)$ est local si et seulement si $n = p^k$ avec p premier et $k \in \mathbb{N}^*$.

2) Soit p un nombre premier. On note :

$$\mathbb{Z}_p = \left\{\frac{a}{b} \;\middle|\; a \in \mathbb{Z},\, b \in \mathbb{N}^*,\, b \wedge p = 1\right\}.$$

Montrer que $\mathbb{Z}_p$ est un sous-anneau de $\mathbb{Q}$ et que $\mathbb{Z}_p$ est un anneau local.

Exercice 1.5 **SF1.12**

Déterminer les $x \in \mathbb{Z}$ tels que :

$$x \equiv 2 \pmod 7 \;;\; x \equiv 1 \pmod 8 \;;\; x \equiv 3 \pmod 9 .$$

Exercice 1.6 **SF1.11**

Soient a, b, c trois entiers. Montrer que si $a^3 + b^3 + c^3$ est multiple de 7, il en est de même du produit abc.

Exercice 1.7 **SF1.13**

Soit $a \in \mathbb{N}^*$. Montrer que $a^{13} - a$ est divisible par 546.

Exercice 1.8 **SF1.10** **SF1.13**

Montrer que pour tout $n \in \mathbb{N}$, $2^{2^{6n+2}} + 3$ est divisible par 19.

Exercice 1.9 **SF1.11**

Théorème de Wilson

1) Soit p un nombre premier. En considérant les solutions de l'équation $x^2 = 1$ dans $\mathbb{Z}/p\mathbb{Z}$, montrer que : $(p-1)! \equiv -1 \pmod{p}$.

2) Montrer que si $n \geqslant 2$ est tel que $(n-1)! \equiv -1 \pmod{n}$, alors n est premier.

Exercice 1.10 **SF1.18** **SF1.19**

1) Soient A et B deux polynômes non constants et premiers entre eux. Montrer qu'il existe un unique couple de polynômes U_0 et V_0 tels que :

$$AU_0 + BV_0 = 1 \ ; \ \deg U_0 < \deg B \ ; \ \deg V_0 < \deg A \, .$$

Exprimer alors en fonction de U_0 et V_0 l'ensemble des couples (U,V) de polynômes tels que $AU + BV = 1$.

2) Déterminer U_0 et V_0 dans le cas : $A = (1-X)^n$ et $B = X^n$ $(n \in \mathbb{N}^*)$.

Exercice 1.11 **SF1.23**

Déterminer les polynômes $P \in \mathbb{C}[X]$ tels que $P(X^2) = P(X)P(X-1)$.

Exercice 1.12 **SF1.25** **SF1.26** **SF1.27**

Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme scindé de degré $n \geqslant 2$ à coefficients réels. On suppose de plus que toutes les racines de P sont simples.

1) Montrer que P' est scindé dans $\mathbb{R}[X]$ et n'admet que des racines simples.

2) Montrer que : $\forall\, x \in \mathbb{R}$, $P'^2(x) - P(x)P''(x) > 0$.

3) Montrer que : $\forall\, k \in [\![1\,;n-1]\!]$, $a_{k-1}a_{k+1} < \dfrac{k}{k+1}a_k^2$.

Exercice 1.13 **SF1.24**

Soit $\alpha \in \mathbb{R}$ et $n \in \mathbb{N}^*$. Déterminer les racines du polynôme :

$$A_n = (X+1)^n - \mathrm{e}^{2\mathrm{i}n\alpha} \, .$$

En déduire la valeur du produit : $\prod_{k=0}^{n-1} \sin\left(\alpha + \dfrac{k\pi}{n}\right)$.

Exercice 1.14 **SF1.21**

Montrer que le polynôme $P = X^4 - 5X^2 + 1$ est irréductible dans $\mathbb{Q}[X]$.

Thème 2 - Espaces vectoriels et applications linéaires

Ce thème est pour l'essentiel consacré à des révisions de 1re année, qui sont indispensables pour la suite.

■ Comment montrer qu'un ensemble est un espace vectoriel ?

SF2.1 Revenir à la définition

On utilisera la définition complète d'un espace vectoriel **[S2.1]** exclusivement dans le cas où il s'agit d'un ensemble ou de lois qui ne sont pas connus et qui viennent d'être définis par l'énoncé.

Exemple

Soit E un $\mathbb{R}$-espace vectoriel.
On munit le produit cartésien $E \times E$ de l'addition usuelle définie par :

$$(x,y) + (x',y') = (x + x',y + y')$$

et de la multiplication externe par un nombre complexe définie par :

$$\forall\,(a,b) \in \mathbb{R}^2,\ \forall\,(x,y) \in E \times E,\ (a + \mathrm{i}b).(x,y) = (ax - by,ay + bx)\,.$$

Montrer que $E \times E$ est alors un $\mathbb{C}$-espace vectoriel (appelé *complexifié* de E).

- *Déjà, on sait que $E \times E$ muni des lois usuelles est un $\mathbb{R}$-espace vectoriel, c'est l'espace vectoriel produit (***[S2.2]***).*
  Cela permet d'affirmer directement que $(E \times E,+)$ est un groupe abélien.
- *On vérifie alors les quatre propriétés de la loi externe.*
  - *Pour tout $(x,y) \in E^2$, $1_{\mathbb{C}}.(x,y) = (1 + 0\mathrm{i}).(x,y) = (x,y)$.*
  - *Soit $\lambda = a + \mathrm{i}b \in \mathbb{C}$ $\big((a,b) \in \mathbb{R}^2\big)$ et (x,y), $(x',y') \in E^2$. On a alors :*

$$\begin{aligned}\lambda.\big[(x,y) + (x',y')\big] &= (a + \mathrm{i}b).(x + x',y + y')\\ &= (a(x + x') - b(y + y'),a(y + y') + b(x + x'))\\ &= (ax - by,ay + bx) + (ax' - by',ay' + bx')\\ &= \lambda.(x,y) + \lambda.(x',y')\,.\end{aligned}$$

  - *Soient $\lambda = a + \mathrm{i}b$ et $\mu = c + \mathrm{i}d$ dans $\mathbb{C}$ et $(x,y) \in E^2$. Alors :*

$$\begin{aligned}(\lambda + \mu).(x,y) &= ((a + c) + \mathrm{i}(b + d)).(x,y)\\ &= ((a + c)x - (b + d)y,(a + c)y + (b + d)x)\\ &= (ax - by,ay + bx) + (cx - dy,cy + dx)\\ &= \lambda.(x,y) + \mu.(x,y)\,.\end{aligned}$$

  - *Soient $\lambda = a + \mathrm{i}b$ et $\mu = c + \mathrm{i}d$ dans $\mathbb{C}$ et $(x,y) \in E^2$. Alors :*

$$\begin{aligned}\lambda.(\mu.(x,y)) &= (a + \mathrm{i}b).(cx - dy,cy + dx)\\ &= (a(cx - dy) - b(cy + dx),a(cy + dx) + b(cx - dy))\end{aligned}$$

$$\begin{aligned} \textit{et :}\ \ (\lambda\mu).(x,y) &= ((ac-bd)+\mathrm{i}(ad+bc)).(x,y) \\ &= ((ac-bd)x-(ad+bc)y,(ac-bd)y+(ad+bc)x) \end{aligned}$$
$$\textit{d'où}\ \ \lambda.(\mu.(x,y)) = (\lambda\mu).(x,y)\,.$$

SF2.2 Utiliser la caractérisation d'un sous-espace vectoriel

Pour montrer qu'un ensemble F muni de lois connues est un espace vectoriel, la méthode générale consiste à montrer que c'est une partie d'un espace vectoriel connu E, puis à vérifier la caractérisation **[S2.3]**.

✓ Il ne faut pas se contenter de montrer la stabilité par combinaisons linéaires, mais aussi préciser, si cela n'est pas dit par l'énoncé : « Montrons que F est un sous-espace vectoriel de ... », et justifier cette inclusion.

✓ Pour vérifier que F est non vide, il suffit de vérifier qu'il contient le vecteur nul de E.

Exemple

Montrer que l'ensemble des applications lipschitziennes de $\mathbb{R}$ dans $\mathbb{R}$ est un espace vectoriel pour les lois usuelles.

Rappelons qu'une application $f\colon \mathbb{R}\to\mathbb{R}$ *est dite* lipschitzienne de rapport $k\in\mathbb{R}_+$ *si :* $\forall\,(x,y)\in\mathbb{R}^2,\ |f(x)-f(y)|\leqslant k\,|x-y|\quad(*)$

Notons L l'ensemble des applications lipschitziennes de $\mathbb{R}$ dans $\mathbb{R}$. Nous allons montrer qu'il s'agit d'un sous-espace vectoriel de l'espace vectoriel $\mathcal{A}(\mathbb{R},\mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$. Pour cela, on vérifie point par point la caractérisation d'un sous-espace vectoriel.

- *L est bien inclus dans $\mathcal{A}(\mathbb{R},\mathbb{R})$, par définition.*
- *L est non vide : en effet, l'application nulle vérifie la relation $(*)$ avec par exemple $k=0$ donc appartient à L ;*
- *si f et g sont lipschitziennes de rapports respectifs k et k' et si λ est un réel, alors pour tout $(x,y)\in\mathbb{R}^2$, on a, grâce à l'inégalité triangulaire :*

$$\begin{aligned} |(\lambda f+g)(x)-(\lambda f+g)(y)| &\leqslant |\lambda|\,|f(x)-f(y)|+|g(x)-g(y)| \\ &\leqslant (|\lambda|\,k+k')\,|x-y|\,. \end{aligned}$$

Ainsi, $\lambda f+g$ est lipschitzienne de rapport $|\lambda|\,k+k'$ donc appartient à L.

On a donc prouvé que L est un sous-espace vectoriel de $\mathcal{A}(\mathbb{R},\mathbb{R})$.

SF2.3 Utiliser la caractérisation d'un sous-espace vectoriel engendré

D'après **[S2.5]**, l'ensemble des combinaisons linéaires d'une famille quelconque de vecteurs d'un espace vectoriel E est un sous-espace vectoriel de E. Ce résultat permet non seulement de démontrer qu'un ensemble est un sous-espace vectoriel, mais aussi d'en trouver directement une famille génératrice.

Exemple

Montrer que l'ensemble F des matrices de la forme $M(a,b,c)=\begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix}$ avec $(a,b,c)\in\mathbb{R}^3$ est un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$.

On a immédiatement : $M(a,b,c) = aI + bJ + cK$ avec $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et $K = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Ainsi F est exactement l'ensemble des combinaisons linéaires des trois matrices I, J et K ; c'est donc le sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$ engendré par ces trois matrices.

✓ *On a ainsi prouvé que F est un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$, et aussi trouvé un système générateur de F. Or il est facile de vérifier que le système (I,J,K) est libre (car $\alpha I + \beta J + \gamma K = 0_3$ implique $\alpha = \beta = \gamma = 0$) ; on a donc en plus déterminé une base de F.*

Exemple

Dans le $\mathbb{R}$-espace vectoriel E des applications de $\mathbb{R}$ dans $\mathbb{R}$ on considère l'ensemble F des applications de la forme :

$$x \mapsto P(x)\cos x + Q(x)\sin x\,,$$

où P et Q décrivent $\mathbb{R}_n[X]$ $(n \in \mathbb{N})$.

Montrer que F est un sous-espace vectoriel de dimension finie de E et en donner une base.

- *F est l'ensemble des applications de la forme :*

$$x \mapsto \sum_{k=0}^{n} a_k x^k \cos x + \sum_{k=0}^{n} b_k x^k \sin x$$

pour $a_0,\dots,a_n$ et $b_0,\dots,b_n$ décrivant $\mathbb{R}$.
C'est donc l'ensemble des combinaisons linéaires des fonctions $f_k : x \mapsto x^k \cos x$ et $g_k : x \mapsto x^k \sin x$ avec $k \in [\![0\,;n]\!]$, c'est-à-dire le sous-espace vectoriel engendré par ces $2n+2$ applications.

En particulier, c'est un sous-espace vectoriel de $\mathcal{A}(\mathbb{R},\mathbb{R})$.

- *Montrons que la famille $(f_0,\dots,f_n,g_0,\dots,g_n)$ est libre.*

S'il existe des réels a_k et b_k tels que $\sum_{k=0}^{n} a_k f_k + \sum_{k=0}^{n} b_k g_k = 0$, alors, pour tout x réel on a :

$$\sum_{k=0}^{n} a_k x^k \cos x + \sum_{k=0}^{n} b_k x^k \sin x = 0\,.$$

Pour tout $x \in 2\pi\mathbb{Z}$ on a alors $\sum_{k=0}^{n} a_k x^k = 0$, c'est-à-dire que le polynôme $\sum_{k=0}^{n} a_k X^k$ possède une infinité de racines. C'est donc le polynôme nul et par suite tous les a_k sont nuls.

En prenant $x \in \frac{\pi}{2} + 2\pi\mathbb{Z}$, on montre de la même façon que tous les b_k sont nuls.

Conclusion : la famille $(f_0,\dots,f_n,g_0,\dots,g_n)$ est libre ; étant génératrice de F, c'en est une base, et $\dim F = 2n+2$.

✓ Cette caractérisation d'un sous-espace vectoriel comme sous-espace vectoriel engendré par une partie est assez fréquente, et permet de simplifier les calculs ; il faut donc apprendre à la repérer et à l'utiliser.

SF2.4 Montrer que c'est le noyau ou l'image d'une application linéaire

On utilise ici les propriétés rappelées dans **[S2.17]**. Là encore, il faut apprendre à repérer cette caractérisation.

Exemple

Soient E un $\mathbb{K}$-espace vectoriel et u, v deux endomorphismes de E.
Montrer que l'ensemble $F = \{x \in E \mid u(x) = v(x)\}$ est un sous-espace vectoriel de E.

Puisque :

$$u(x) = v(x) \iff (u - v)(x) = 0\,,$$

F est simplement le noyau de l'endomorphisme $u - v$ (qui est bien un endomorphisme comme différence de deux endomorphismes). F est donc un sous-espace vectoriel de E.

SF2.5 Montrer qu'il s'agit de l'intersection de sous-espaces vectoriels

Décrire un ensemble comme intersection de sous-espaces vectoriels plus simples est une situation fréquente. La reconnaître permet d'alléger les raisonnements.

Exemple

Soit E l'espace vectoriel $\mathscr{C}^\infty(\mathbb{R},\mathbb{R})$, n un entier naturel et F l'ensemble des applications $f \in E$ telles que, pour tout $k \in [\![0\,;n]\!]$, la dérivée k-ième de f s'annule en 0.

Montrer que F est un sous-espace vectoriel de E.

Pour tout entier k l'application $\varphi_k : f \mapsto f^{(k)}(0)$ est une forme linéaire sur E (comme composée k fois de l'opérateur de dérivation et de la forme linéaire $f \mapsto f(0)$).
L'ensemble des $f \in E$ telles que $f^{(k)}(0) = 0$ est le noyau de φ_k ; c'est donc un sous-espace vectoriel de E ; puis $F = \bigcap_{k=0}^{n} \operatorname{Ker} \varphi_k$ est l'intersection de ces sous-espaces vectoriels, c'est donc aussi un sous-espace vectoriel de E.

Exemple

Montrer que l'ensemble

$$F = \left\{(x,y,z,t) \in \mathbb{R}^4 \;\middle|\; x + y + z + t = x + 2y + 3z + 4t = 0\right\}$$

est un sous-espace vectoriel de $\mathbb{R}^4$.

L'application $\varphi : (x,y,z,t) \mapsto x + y + z + t$ est une forme linéaire sur $\mathbb{R}^4$.
L'ensemble $H_1 = \left\{(x,y,z,t) \in \mathbb{R}^4 \;\middle|\; x + y + z + t = 0\right\}$ est le noyau de cette forme linéaire, c'est donc un sous-espace vectoriel de $\mathbb{R}^4$.

On démontre de la même façon que $H_2 = \{(x,y,z,t) \mid x + 2y + 3z + 4t = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^4$.

Ainsi, $F = H_1 \cap H_2$ est bien un sous-espace vectoriel de $\mathbb{R}^4$.

✓ *Le fait que l'application φ ci-dessus est une forme linéaire fait partie du cours (**[S2.20]**) et n'a pas à être redémontré. On peut d'ailleurs aussi dire, toujours d'après le cours, que H_1 et H_2 sont des hyperplans de $\mathbb{R}^4$.*

✓ Tous les exemples précédents montrent qu'il est très important de bien connaître un certain nombre d'espaces vectoriels usuels et d'applications linéaires usuelles.

■ Comment étudier l'indépendance linéaire d'une famille de vecteurs ?

SF2.6 Revenir à la définition

Dans beaucoup de cas, pour montrer qu'une famille finie de vecteurs est libre, on revient simplement à la définition **[S2.7]** : on écrit qu'une combinaison linéaire de ces vecteurs est égale au vecteur nul, et on démontre que les coefficients sont nuls. Nous avons d'ailleurs utilisé cette méthode dans des exemples précédents.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et $(e_1, \dots, e_n)$ une base de E. Soient $\lambda_1, \dots, \lambda_n$ des scalaires, et $u = \sum_{i=1}^{n} \lambda_i e_i$. Pour tout $i \in [\![1\,;n]\!]$ on pose $v_i = u + e_i$. Déterminer une condition nécessaire et suffisante portant sur les λ_i pour que la famille $(v_1, \dots, v_n)$ soit libre.

Supposons qu'il existe des scalaires $(\alpha_1, \dots, \alpha_n)$ tels que $\sum_{k=1}^{n} \alpha_k v_k = 0$. On obtient alors, en remplaçant les v_k par leurs valeurs :

$$\left(\sum_{k=1}^{n} \alpha_k\right)\left(\sum_{i=1}^{n} \lambda_i e_i\right) + \sum_{k=1}^{n} \alpha_k e_k = 0\,.$$

En considérant alors les coordonnées sur chaque vecteur e_i on trouve :

$$\forall\, i \in [\![1\,;n]\!],\ \left(\sum_{k=1}^{n} \alpha_k\right)\lambda_i + \alpha_i = 0\,.$$

Pour simplifier, notons $s = \sum_{k=1}^{n} \alpha_k$. On a donc un système de n équations :

$$\forall\, i \in [\![1\,;n]\!],\ s\lambda_i + \alpha_i = 0 \quad (*)$$

En additionnant ces n équations on obtient : $s\left(\sum_{i=1}^{n} \lambda_i\right) + s = 0$.

Deux cas se présentent donc.

– *Si $\sum_{i=1}^{n} \lambda_i + 1 \neq 0$, on obtient $s = 0$ puis en remplaçant dans $(*)$ on en tire $\alpha_i = 0$ pour tout i ; donc la famille $(v_1, \dots, v_n)$ est libre.*

– *Si* $\sum_{i=1}^{n} \lambda_i = -1$, *on vérifie facilement que* $\sum_{i=1}^{n} \lambda_i v_i = 0$, *donc, puisque les* λ_i *ne sont pas tous nuls car leur somme vaut* -1, *on en déduit que la famille* $(v_1, \dots, v_n)$ *est liée.*

Conclusion : la famille $(v_1, \dots, v_n)$ *est libre si et seulement si* $\sum_{i=1}^{n} \lambda_i + 1 \neq 0$.

✓ Cet exemple montre que, lorsque l'on travaille dans un espace vectoriel de dimension finie, il est indispensable de savoir résoudre proprement et rigoureusement un système d'équations linéaires, en particulier lorsque celui-ci comporte un paramètre.

SF2.7 Raisonner par l'absurde

Il s'agit ici encore d'utiliser la définition, mais en raisonnant par l'absurde : on suppose qu'il existe une combinaison linéaire non triviale, et l'on essaye d'aboutir à une contradiction.

Exemple

Démontrer que la famille des applications $f_a : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \mathrm{e}^{ax} \end{cases}$ lorsque a décrit $\mathbb{R}$ est libre.

Rappelons déjà que, pour montrer qu'une famille infinie de vecteurs est libre, il suffit de montrer que toute sous-famille finie l'est.
Considérons donc une famille finie $f_{a_1}, f_{a_2}, \dots, f_{a_n}$, *où les* a_i *sont des réels distincts, et supposons par l'absurde qu'il existe des scalaires* $\lambda_1, \dots, \lambda_n$ *non tous nuls tels que* $\sum_{i=1}^{n} \lambda_i f_{a_i} = 0$.

Quitte à changer l'ordre, on peut bien sûr supposer $a_1 < a_2 < \dots < a_n$; *on peut ensuite supposer* $\lambda_n \neq 0$ *(ce n'est pas la peine d'écrire des termes nuls !). On aurait donc :*

$$\forall x \in \mathbb{R},\ \lambda_n \mathrm{e}^{a_n x} + \lambda_{n-1} \mathrm{e}^{a_{n-1} x} + \dots + \lambda_1 \mathrm{e}^{a_1 x} = 0 \quad (*)$$

Mais pour tout $i \in [\![1 ; n-1]\!]$, $\mathrm{e}^{a_i x}$ *est négligeable devant* $\mathrm{e}^{a_n x}$ *lorsque* $x \to +\infty$ *puisque* $a_i < a_n$.
Donc $\lambda_n \mathrm{e}^{a_n x} + \lambda_{n-1} \mathrm{e}^{a_{n-1} x} + \dots + \lambda_1 \mathrm{e}^{a_1 x} \underset{x \to +\infty}{\sim} \lambda_n \mathrm{e}^{a_n x}$, *ce qui contredit* $(*)$ *et achève la démonstration.*

SF2.8 Utiliser une récurrence

Là encore, il s'agit d'utiliser la définition. Pour montrer que les coefficients de la combinaison linéaire sont nuls, on peut faire une récurrence sur le cardinal de la famille.

Exemple

Pour $k \in \mathbb{N}^*$ et $x \in \mathbb{R}$, on pose $f_k(x) = \sin kx$.

Montrer que, pour tout $n \in \mathbb{N}^*$, la famille $(f_1, \dots, f_n)$ est libre.

On procède par récurrence sur $n \in \mathbb{N}^*$.

– *Pour $n = 1$ la famille est formée d'un seul élément non nul (la fonction $f_1 : x \mapsto \sin x$) donc est libre.*

– *Supposons démontré le résultat à l'ordre $n-1$ ($n \geqslant 2$), et supposons qu'il existe des scalaires $(\lambda_1, \ldots, \lambda_n)$ tels que $\sum_{k=1}^{n} \lambda_k f_k = 0$ (1).*

En dérivant 2 fois cette relation, on obtient : $\sum_{k=1}^{n} k^2 \lambda_k f_k = 0$ (2).

En effectuant alors l'opération $(2) - n^2 \times (1)$, les termes en f_n s'éliminent et il reste :

$$\sum_{k=1}^{n-1} (k^2 - n^2) \lambda_k f_k = 0\,.$$

D'après l'hypothèse de récurrence, la famille $(f_1, \ldots, f_{n-1})$ est libre donc :

$$\forall\, k \in [\![1\,;n-1]\!],\ (k^2 - n^2)\lambda_k = 0\,.$$

D'où $\lambda_k = 0$ pour tout $k \in [\![1\,;n-1]\!]$, puis ensuite $\lambda_n = 0$ d'après (1). Finalement, tous les λ_k sont nuls ; cela prouve que la famille $(f_1, \ldots, f_n)$ est libre, ce qui est la propriété voulue à l'ordre n et achève la récurrence.

SF2.9 Utiliser directement des résultats du cours

Il y a dans le cours certaines familles libres et bases à connaître. Dans certains cas, il suffit d'utiliser directement ce résultat. Voici quelques situations fréquentes.

1. Polynômes de degrés distincts

D'après le cours, toute famille finie de polynômes non nuls de degrés distincts est libre. Vous pouvez utiliser directement ce résultat, sans le redémontrer.
On obtient ainsi directement l'indépendance linéaire des familles suivantes (utilisées assez fréquemment) :

– les polynômes $(X - a)^k$ pour $k \in \mathbb{N}$;

– les polynômes $1, X, X(X-1), X(X-1)(X-2), \ldots$

2. Polynômes de valuations distinctes

Rappelons que la *valuation* $v(P)$ d'un polynôme $P \in \mathbb{K}[X]$ est le plus grand entier i tel que X^i divise P (si P est le polynôme nul, on la pose égale à $+\infty$). Il s'agit donc du degré du terme de plus petit degré de P.

Démontrons, par récurrence sur n, que toute famille de polynômes $(P_k)_{0 \leqslant k \leqslant n}$ tels que $v(P_k) = k$ pour tout $k \in [\![0\,;n]\!]$ est libre.

– *Le résultat est immédiat lorsque $n = 0$.*

– *Supposons le résultat acquis à l'ordre $n-1$ ($n \geqslant 1$), et soit $(P_k)_{0 \leqslant k \leqslant n}$ une famille de polynômes tels que $v(P_k) = k$ pour tout k.*

Supposons qu'il existe des scalaires $(\lambda_0, \ldots, \lambda_n)$ tels que $\sum_{k=0}^{n} \lambda_k P_k = 0$.

On aura alors $\sum_{k=0}^{n} \lambda_k P_k(0) = 0$. Mais pour $k \geqslant 1$ P_k est divisible par X donc $P_k(0) = 0$; l'égalité précédente se réduit donc à $\lambda_0 P_0(0) = 0$ et puisque P_0 est de valuation nulle, on a $P_0(0) \neq 0$ d'où $\lambda_0 = 0$.

On a donc $\sum_{k=1}^{n} \lambda_k P_k = 0$ *d'où en divisant par* X *:* $\sum_{k=1}^{n} \lambda_k \frac{P_k}{X} = 0$ *ou* $\sum_{k=0}^{n-1} \lambda_{k+1} \frac{P_{k+1}}{X} = 0$. *Les polynômes* $\frac{P_{k+1}}{X}$ *pour* $0 \leqslant k \leqslant n-1$ *forment une famille de polynômes tels que* $v\left(\frac{P_{k+1}}{X}\right) = k$ *pour tout* $k \in [\![0\,;n-1]\!]$ *; on peut donc leur appliquer l'hypothèse de récurrence, d'où* $\lambda_{k+1} = 0$ *pour* $0 \leqslant k \leqslant n-1$.

Finalement tous les coefficients λ_k *pour* $0 \leqslant k \leqslant n$ *sont nuls, ce qui prouve que la famille* $(P_k)_{0\leqslant k\leqslant n}$ *est libre et achève la récurrence.*

Il en résulte que toute famille de polynômes non nuls de valuations distinctes est libre, puisque toute sous-famille finie sera incluse dans une famille du type précédent.

✓ Comme application classique de ce résultat, on obtient par exemple que la famille des polynômes $X^k(1-X)^{n-k}$ pour $k \in [\![0\,;n]\!]$ est une base de $\mathbb{R}_n[X]$.

3. Famille de vecteurs propres

D'après le cours, toute famille de vecteurs propres d'un endomorphisme, associés à des valeurs propres distinctes, est libre. Cela permet dans certains cas de démontrer très rapidement l'indépendance linéaire de certaines familles.

Exemple

Démontrer que la famille des applications $f_a : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \mathrm{e}^{ax} \end{cases}$ lorsque a décrit $\mathbb{R}$ est libre.

Nous avons déjà traité cet exemple, mais la démonstration qui suit est plus rapide.
Considérons l'espace vectoriel $E = \mathscr{C}^\infty(\mathbb{R},\mathbb{R})$ *des applications de classe* $\mathscr{C}^\infty$ *de* $\mathbb{R}$ *dans* $\mathbb{R}$. *Les applications* f_a *sont des vecteurs de* E.

L'application $D : f \in E \longmapsto f'$ *est clairement un endomorphisme de* E. *Pour tout* $a \in \mathbb{R}$, $D(f_a) = af_a$ *donc* f_a *est un vecteur propre de* D *associé à la valeur propre* a. *En vertu du résultat de cours rappelé auparavant, on peut donc conclure directement qu'il s'agit d'une famille libre.*

4. Utilisation de bases ou de familles libres connues

Si une famille est une sous-famille d'une base ou d'une famille libre connue, elle est libre d'après les propriétés rappelées en **[S2.7]**.

Exemple

Démontrer que l'ensemble $\mathcal{T}_n^+(\mathbb{K})$ des matrices triangulaires supérieures d'ordre n est un espace vectoriel et en préciser la dimension.

Notons $(E_{ij})_{1\leqslant i,j\leqslant n}$ *la base canonique de* $\mathcal{M}_n(\mathbb{K})$. *Toute matrice* $A = (a_{ij})$ *de* $\mathcal{M}_n(\mathbb{K})$ *s'écrit* $A = \sum_{1\leqslant i,j\leqslant n} a_{ij}E_{ij}$ *et toute matrice triangulaire supérieure* $T = (a_{ij})$ *s'écrit donc, puisque* $a_{ij} = 0$ *dès que* $i > j$ *:*

$$T = \sum_{1\leqslant i\leqslant j\leqslant n} a_{ij}E_{ij}\,.$$

Cela montre que $\mathcal{T}_n^+(\mathbb{K})$ est le sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ engendré par les matrices E_{ij} pour $1 \leqslant i \leqslant j \leqslant n$. Puisque cet ensemble de matrices est une partie de la base canonique, elles forment un système libre, donc une base de $\mathcal{T}_n^+(\mathbb{K})$.

Il en résulte que la dimension de $\mathcal{T}_n^+(\mathbb{K})$ est égale au nombre de couples $(i,j) \in [\![1\,;n]\!]^2$ tels que $i \leqslant j$, c'est-à-dire $\dfrac{n(n+1)}{2}$.

■ Comment trouver le rang d'une famille de vecteurs ?

SF2.10 Trouver une sous-famille libre de cardinal maximum

Cette méthode demande un peu d'intuition, pour arriver à repérer dans une famille donnée des vecteurs déjà linéairement indépendants.

Exemple

Dans l'espace vectoriel E des applications de $\mathbb{R}$ dans $\mathbb{R}$, trouver le rang de la famille $(f_1,\dots,f_6)$ où pour tout réel x :

$$f_1(x) = \sin x \qquad f_2(x) = \cos x \qquad f_3(x) = \cos 2x \qquad f_4(x) = \sin(x+1)$$
$$f_5(x) = \sin^2 x \qquad f_6(x) = \cos^2 x\,.$$

- *Montrons déjà que la famille (f_1,f_2,f_5,f_6) est libre : s'il existe des réels λ_i tels que $\lambda_1 f_1 + \lambda_2 f_2 + \lambda_5 f_5 + \lambda_6 f_6 = 0$ alors on a :*
$$\forall\, x \in \mathbb{R},\ \lambda_1 \sin x + \lambda_2 \cos x + \lambda_5 \sin^2 x + \lambda_6 \cos^2 x = 0\,.$$
On en déduit alors facilement, en faisant par exemple $x = 0$ et $x = \pi$, que $\lambda_2 = \lambda_6 = 0$, puis en faisant $x = \dfrac{\pi}{2}$ et $x = -\dfrac{\pi}{2}$ que $\lambda_1 = \lambda_5 = 0$, ce qui démontre le résultat voulu.
- *D'après les formules trigonométriques : $\sin(x+1) = \sin 1 \cos x + \sin x \cos 1$ donc $f_4 = \cos 1 f_1 + \sin 1 f_2$ est combinaison linéaire de f_1 et f_2.*
- *On a aussi : $\cos 2x = \cos^2 x - \sin^2 x$ donc $f_3 = f_6 - f_5$ est combinaison linéaire de f_5 et f_6.*

En conclusion, $\mathrm{rg}(f_1,f_2,f_3,f_4,f_5,f_6) = \mathrm{rg}(f_1,f_2,f_5,f_6) = 4$.

SF2.11 Utiliser la méthode du pivot

Nous ne revenons pas en détail sur cette méthode, que vous avez étudiée en première année. Nous donnons simplement un exemple, en vous suggérant de faire les calculs par vous-même puis de comparer avec la solution donnée. Vous remarquerez que, compte tenu de la présence d'un paramètre, nous avons choisi les opérations à effectuer de façon à ne pas avoir à discuter selon la valeur de ce paramètre lors des calculs, donc en choisissant toujours un pivot qui est une constante, quitte à faire des permutations supplémentaires sur les lignes ou colonnes.

Exemple

Soit m un paramètre réel. Déterminer selon la valeur de m le rang de la famille de vecteurs de $\mathbb{R}^4$:

$$a = (2,1,m,2m)\,,\ b = (1,0,m,3)\,,\ c = (1,-1,1,m) \ \text{ et } \ d = (3,0,1,4)\,.$$

La matrice de la famille (b,a,c,d) dans la base canonique de $\mathbb{R}^4$ est :

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -1 & 0 \\ m & m & 1 & 1 \\ 3 & 2m & m & 4 \end{pmatrix}.$$

(écrire les vecteurs dans cet ordre permet déjà d'avoir un coefficient 1 en haut à gauche).
Par les opérations élémentaires $L_3 \leftarrow L_3 - mL_1$ et $L_4 \leftarrow L_4 - 3L_1$, le rang de A est aussi celui de la matrice :

$$A_2 = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & -m & 1-m & 1-3m \\ 0 & 2m-6 & m-3 & -5 \end{pmatrix}.$$

Par les opérations élémentaires $L_3 \leftarrow L_3 + mL_2$ et $L_4 \leftarrow L_4 - (2m-6)L_2$, le rang de A est aussi celui de la matrice :

$$A_3 = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1-2m & 1-3m \\ 0 & 0 & 3m-9 & -5 \end{pmatrix},$$

donc, en échangeant les deux dernières colonnes et les deux dernières lignes, c'est aussi celui de la matrice :

$$A_4 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -5 & 3m-9 \\ 0 & 0 & 1-3m & 1-2m \end{pmatrix},$$

et, enfin, en faisant l'opération $L_4 \leftarrow L_4 + \dfrac{1-3m}{5} L_3$, le rang de la matrice initiale est aussi celui de la matrice :

$$A_4 = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -5 & 3m-9 \\ 0 & 0 & 0 & \dfrac{-9m^2+20m-4}{5} \end{pmatrix}.$$

On en déduit que :

- *si $-9m^2 + 20m - 4 = 0$ c'est-à-dire si $m \in \left\{2, \dfrac{2}{9}\right\}$, le rang de la famille est égal à 3 ;*
- *sinon, le rang de la famille est égal à 4 (et c'est alors une base de $\mathbb{R}^4$).*

SF2.12 Utiliser le déterminant

Cette méthode n'est évidemment utilisable que si le nombre de vecteurs est égal à la dimension de l'espace.

Exemple

Dans $\mathbb{R}^4$ déterminer selon la valeur de m le rang du système de vecteurs

$$a = (m,1,1,1)\,,\ b = (1,m,1,1)\,,\ c = (1,1,m,1) \ \text{ et } \ d = (1,1,1,m)\,.$$

La matrice de ces vecteurs dans la base canonique de $\mathbb{R}^4$ est :

$$A = \begin{pmatrix} m & 1 & 1 & 1 \\ 1 & m & 1 & 1 \\ 1 & 1 & m & 1 \\ 1 & 1 & 1 & m \end{pmatrix}.$$

Un calcul assez simple, que nous ne reproduisons pas ici mais que nous vous conseillons fortement de faire (vous devez *y arriver, en un temps raisonnable), donne :* $\det A = (m-1)^3(m+3)$ *(voir aussi l'exemple de* **[SF3.36]** *page 280). On en déduit :*

- *si* $m \notin \{-3,1\}$, $\det A \neq 0$ *donc la famille* (a,b,c,d) *est une base de* $\mathbb{R}^4$ *; elle est donc de rang* 4.
- *si* $m = 1$, *on a alors* $a = b = c = d$ *donc la famille est de rang* 1.
- *si* $m = -3$, *le déterminant de* A *étant nul,* $\operatorname{rg} A$ *est inférieur ou égal à* 3. *On calcule alors* $\begin{vmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -3 \end{vmatrix} = -16$ *; ce déterminant étant non nul, les trois premières colonnes de la matrice* A *sont linéairement indépendantes, donc la matrice* A *est de rang* 3, *qui est aussi le rang de la famille* (a,b,c,d).

■ Comment montrer qu'une famille de vecteurs est une base ?

SF2.13 Montrer qu'elle est libre et génératrice

Pour montrer qu'une famille est une base d'un espace vectoriel E, on peut vérifier séparément qu'elle libre et génératrice de E. Il s'agit de la méthode la plus simple dans son principe, mais pas toujours facile à mettre en œuvre. On la réservera au cas où E est de dimension infinie, ou bien lorsque l'on ne connaît pas sa dimension ; sinon, on préfèrera les méthodes qui utilisent la dimension, exposées plus loin.

Exemple

Soit $(P_n)_{n\in\mathbb{N}}$ une famille de polynômes de $\mathbb{K}[X]$ telle que $\deg(P_n) = n$ pour tout entier n.
Montrer qu'il s'agit d'une base de $\mathbb{K}[X]$.

- *Les polynômes* P_n *étant non nuls et de degrés distincts forment une famille libre (cf.* **[SF2.9]***).*
- *Soit* Q *un polynôme quelconque de* $\mathbb{K}[X]$. *Il existe alors un entier* N *tel que* Q *soit élément de* $\mathbb{K}_N[X]$ *(il suffit de prendre un entier* N *supérieur au degré de* Q*).*

  Puisque la famille $(P_0,\dots,P_N)$ *est une famille libre formée de* $N+1$ *éléments de l'espace vectoriel* $\mathbb{K}_N[X]$ *qui est de dimension* $N+1$, *c'en est une base, donc il existe des scalaires* λ_i *tels que* $Q = \sum_{i=0}^{N} \lambda_i P_i$.

  Q *est donc combinaison linéaire (finie) de la famille* $(P_n)_{n\in\mathbb{N}}$.

Cela montre que la famille $(P_n)_{n\in\mathbb{N}}$ *est une famille génératrice de* $\mathbb{K}[X]$ *et, finalement, c'en est une base.*

Exemple

Soit $I = [a\,;b]$ un segment de $\mathbb{R}$ $(a < b)$, et $(x_0,x_1,\ldots,x_n)$ une subdivision donnée de ce segment, c'est-à-dire que : $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$.

On note E l'ensemble des applications f de I dans $\mathbb{R}$ telles que la restriction de f à chaque intervalle $]x_i\,;x_{i+1}[$ pour $0 \leqslant i \leqslant n-1$ est une fonction affine.

Montrer que E un un $\mathbb{R}$-espace vectoriel de dimension finie et en donner une base.

- *Commençons par démontrer que E est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $\mathcal{A}(I,\mathbb{R})$ des applications de I dans $\mathbb{R}$:*
  - *E est inclus dans $\mathcal{A}(I,\mathbb{R})$ par définition ;*
  - *l'application nulle appartient à E, donc E est non vide ;*
  - *si f et g sont dans E et si $\lambda \in \mathbb{R}$, f et g sont affines sur chaque intervalle $]x_i\,;x_{i+1}[$ donc $\lambda f + g$ l'est aussi, ce qui prouve que $\lambda f + g$ appartient à E.*
- *Soit $f \in E$. f est entièrement déterminée par :*
  - *sa restriction aux intervalles de la subdivision : il existe des réels a_i et b_i pour $0 \leqslant i \leqslant n-1$ tels que pour tout $x \in]x_i\,;x_{i+1}[$ l'on ait $f(x) = a_i x + b_i$.*
  - *et sa valeur aux points de cette subdivision : notons $c_i = f(x_i)$ pour tout $i \in [\![0\,;n]\!]$.*

  Cela suggère d'introduire les applications suivantes :
  - *pour tout $i \in [\![0\,;n-1]\!]$, soit $\alpha_i : x \in [a\,;b] \mapsto \begin{cases} x & \text{si } x \in]x_i\,;x_{i+1}[\\ 0 & \text{sinon} \end{cases}$ et*

    $\beta_i : x \in [a\,;b] \mapsto \begin{cases} 1 & \text{si } x \in]x_i\,;x_{i+1}[\\ 0 & \text{sinon.} \end{cases}$
  - *pour tout $i \in [\![0\,;n]\!]$ soit $\gamma_i : x \in [a\,;b] \mapsto \begin{cases} 1 & \text{si } x = x_i \\ 0 & \text{sinon.} \end{cases}$*

  Ces applications sont bien des éléments de E, et l'application f précédente peut s'écrire :
  $$f = \sum_{i=0}^{n-1} a_i\alpha_i + \sum_{i=0}^{n-1} b_i\beta_i + \sum_{i=0}^{n} c_i\gamma_i\,.$$
  Donc la famille des applications α_i, β_i, γ_i est génératrice de E.
- *Cette famille est libre puisque, en reprenant les notations précédentes, si*
  $$\sum_{i=0}^{n-1} a_i\alpha_i + \sum_{i=0}^{n-1} b_i\beta_i + \sum_{i=0}^{n} c_i\gamma_i = 0\,,$$
  alors, en prenant la valeur de cette application en chaque point x_i on obtient $c_i = 0$ puis sur chaque intervalle on aura $a_i x + b_i = 0$ pour tout $x \in]x_i\,;x_{i+1}[$ ce qui entraîne $a_i = b_i = 0$.
- *Par conséquent, la famille formée des applications α_i, β_i pour $i \in [\![0\,;n-1]\!]$ et γ_i pour $i \in [\![0\,;n]\!]$ est une base de E ; E est donc un espace vectoriel de dimension finie, et* $\dim E = n + n + (n+1) = 3n+1$.

SF2.14 Montrer l'existence et l'unicité de la décomposition

Pour montrer qu'une famille $(e_1,\dots,e_n)$ est une base d'un espace vectoriel E, on peut montrer directement que tout vecteur $x \in E$ s'écrit de manière unique comme combinaison linéaire des e_i. Cette méthode est elle aussi assez longue, mais elle a l'avantage de donner directement les coordonnées d'un vecteur dans la nouvelle base. Elle est donc à privilégier si l'on a par la suite besoin de ces coordonnées.

Exemple

Dans $\mathbb{R}^3$, soient $u = (1,1,0)$, $v = (1,0,1)$ et $w = (0,1,1)$. Montrer qu'il s'agit d'une base et donner les coordonnées d'un vecteur $(x,y,z) \in \mathbb{R}^3$ dans cette base.

Nous répondrons directement aux deux questions.
Soit $(x,y,z) \in \mathbb{R}^3$. On cherche des réels α,β,γ tels que :

$$(x,y,z) = \alpha u + \beta v + \gamma w\,.$$

En identifiant les composantes des deux membres, cela équivaut au système :

$$\begin{cases} x = \alpha + \beta & (1) \\ y = \alpha + \gamma & (2) \\ z = \beta + \gamma & (3) \end{cases}.$$

$(1) + (2) - (3)$ donne $2\alpha = x + y - z$ et on a de même $2\beta = x - y + z$ et $2\gamma = -x + y + z$.
Ainsi α, β et γ existent et sont uniques (en fonction de x, y, z) : cela prouve que la famille (u,v,w) est une base de $\mathbb{R}^3$. De plus, par le calcul précédent, on a montré que les coordonnées d'un vecteur (x,y,z) dans cette base sont :

$$\alpha = \frac{x+y-z}{2}\,,\ \beta = \frac{x-y+z}{2} \quad \text{et} \quad \gamma = \frac{-x+y+z}{2}\,.$$

SF2.15 Utiliser le cardinal de la famille et la dimension

On sait (**[S2.11]**) que dans un espace de dimension finie n toute famille libre ayant exactement n éléments et que toute famille génératrice ayant exactement n éléments sont des bases.
Si l'on connaît la dimension de l'espace (ce qui n'est pas toujours le cas !), il s'agit de la méthode à privilégier. On est alors ramené à prouver l'indépendance linéaire ou à déterminer le rang d'une famille de vecteurs, ce qui a fait l'objet des deux problématiques précédentes.

Exemple

Soit $n \in \mathbb{N}$. Démontrer que la famille des polynômes

$$1, X, X(X-1), \dots, X(X-1)\cdots(X-n+1)$$

forme une base de $\mathbb{R}_n[X]$.

- *On remarque déjà que ces polynômes sont bien tous de degré $\leqslant n$ (chaque produit $X(X-1)\cdots(X-k+1)$ comporte k facteurs), donc sont bien éléments de $\mathbb{R}_n[X]$.*
- *Ces polynômes étant tous non nuls et de degrés distincts forment une famille libre.*

- *Cette famille est de cardinal $n+1$, et $n+1 = \dim \mathbb{R}_n[X]$.*

C'est donc une base de $\mathbb{R}_n[X]$.

✓ *Cette base est utilisée pour simplifier certains calculs : il y a un exemple d'utilisation en* **[SF11.13]**. *On peut aussi parfois utiliser la base formée des polynômes $1, X, X(X+1), \dots, X(X+1)\cdots(X+n-1)$.*

SF2.16 Utiliser un isomorphisme

Il s'agit ici d'utiliser la propriété **[S2.18]** : l'image d'une base par un isomorphisme d'espaces vectoriels est une base.

Exemple

Soient $(a_0, \dots, a_n)$ $n+1$ scalaires deux à deux distincts. Pour tout $i \in [\![0\,;n]\!]$, on considère le polynôme $L_i = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{X-a_j}{a_i-a_j}$.

En considérant l'application $\varphi : \left\{ \begin{array}{ccc} \mathbb{K}_n[X] & \longrightarrow & \mathbb{K}^{n+1} \\ P & \longmapsto & \big(P(a_0), P(a_1), \dots, P(a_n)\big) \end{array} \right.$, montrer que la famille $(L_i)_{0\leqslant i\leqslant n}$ est une base de $\mathbb{K}_n[X]$.

- *Montrons d'abord que φ est un isomorphisme d'espaces vectoriels.*
  - *φ est une application linéaire : cela se démontre sans difficulté.*
  - *φ est injective : pour cela, on montre que son noyau est réduit au vecteur nul.*

    Soit $P \in \mathbb{K}_n[X]$ tel que $\varphi(P) = 0_{\mathbb{K}^{n+1}}$. Alors pour tout $i \in [\![0\,;n]\!]$ on a $P(a_i) = 0$; ainsi P, polynôme de degré $\leqslant n$, possède-t-il au moins $n+1$ racines distinctes : c'est donc le polynôme nul.

    ✓ *Le raisonnement ci-dessus prouve seulement, en réalité, l'inclusion $\operatorname{Ker}\varphi \subset \{0\}$; mais l'inclusion inverse étant toujours vraie, on ne la mentionne jamais.*
  - *φ est donc une application linéaire injective entre deux espaces vectoriels de* même *dimension $(n+1)$; d'après* **[S2.19]**, *c'est un isomorphisme.*
- *Les polynômes L_i sont de degré n, ce sont donc bien des éléments de $\mathbb{K}_n[X]$. De plus si $j \neq i$, a_j est par construction racine de L_i donc $L_i(a_j) = 0$; et l'on a aussi évidemment $L_i(a_i) = 1$. Ainsi, l'image de L_i par φ est le vecteur $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ où le 1 est à la $(i+1)^e$ place, c'est-à-dire le $(i+1)^e$ vecteur de la base canonique de $\mathbb{K}^{n+1}$.*

  Ainsi, la famille $(L_i)_{0\leqslant i\leqslant n}$ est l'image réciproque par φ de la base canonique de $\mathbb{K}^{n+1}$; φ étant un isomorphisme, c'est bien une base de $\mathbb{K}_n[X]$.

SF2.17 Utiliser le déterminant

Soit E un espace vectoriel de dimension $n \geqslant 1$, où l'on connaît déjà une base $\mathscr{B}$. Alors une famille ayant exactement n vecteurs est une base si et seulement si son déterminant dans la base $\mathscr{B}$ est non nul. On peut donc utiliser cette caractérisation

lorsque la dimension de l'espace est petite ($\leqslant 4$ en général), ou lorsque le calcul du déterminant est facile. En dimension 3, il s'agit de la méthode à privilégier.

L'exemple suivant a déjà été traité page 218 ; cela illustre le fait qu'il y a souvent plusieurs méthodes pour traiter un problème d'algèbre linéaire.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et $\mathscr{B} = (e_1,\dots,e_n)$ une base de E. Soient $\lambda_1,\dots,\lambda_n \in \mathbb{K}$, et $u = \sum_{i=1}^{n} \lambda_i e_i$. Pour tout $i \in [\![1\,;n]\!]$ on pose $v_i = u + e_i$.

Déterminer une condition nécessaire et suffisante portant sur les λ_i pour que la famille $(v_1,\dots,v_n)$ soit une base de E.

On calcule le déterminant de la famille $(v_1,\dots,v_n)$ dans la base $\mathscr{B}$ en utilisant la multilinéarité :

$$\begin{aligned}\det_{\mathscr{B}}(v_1,\dots,v_n) &= \det_{\mathscr{B}}(u+e_1,u+e_2,\dots,u+e_n)\\ &= \det_{\mathscr{B}}(e_1,\dots,e_n) + \det_{\mathscr{B}}(u,e_2,\dots,e_n)\\ &\quad + \det_{\mathscr{B}}(e_1,u,e_3,\dots,e_n) + \cdots + \det_{\mathscr{B}}(e_1,\dots,e_{n-1},u)\,,\end{aligned}$$

les autres termes du développement étant nuls car ce sont des déterminants de familles comportant au moins deux fois le vecteur u.
En développant chaque déterminant

$$\det_{\mathscr{B}}(e_1,\dots,e_{i-1},u,e_{i+1},\dots,e_n) = \begin{vmatrix} 1 & & & \lambda_1 & & \\ & \ddots & & \vdots & & \\ & & 1 & \lambda_{i-1} & & \\ & & & \lambda_i & & \\ & & & \lambda_{i+1} & 1 & \\ & 0 & & \vdots & & \ddots \\ & & & \lambda_n & & 1 \end{vmatrix}$$

selon la i^e ligne, on voit qu'il est égal à λ_i donc :

$$\det_{\mathscr{B}}(v_1,\dots,v_n) = \det_{\mathscr{B}}(e_1,\dots,e_n) + \sum_{i=1}^{n} \lambda_i = 1 + \sum_{i=1}^{n} \lambda_i\,.$$

On retrouve le résultat obtenu en **[SF2.6]** *page 218 : la famille $(v_1,\dots,v_n)$ est libre (donc ce sera une base de E) si et seulement si $\sum_{i=1}^{n} \lambda_i + 1 \neq 0$.*

■ Comment trouver la dimension d'un espace vectoriel ?

SF2.18 Trouver directement une base

Il s'agit évidemment de la méthode à laquelle on peut penser en premier puisque par définition la dimension d'un espace vectoriel est le cardinal d'une base. Voici un premier exemple élémentaire.

Exemple

Dans $\mathbb{R}^4$, soient $a = (1,2,3,4)$, $b = (1,1,1,3)$, $c = (2,1,1,1)$ et $d = (3,1,0,3)$. Trouver la dimension de $F = \mathrm{Vect}(a,b,c,d)$.

La famille (a,b,c,d) est, par définition, génératrice de F. D'autre part, la matrice de ces 4 vecteurs dans la base canonique de $\mathbb{R}^4$ est égale à :

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 2 & 1 & 1 & 1 \\ 3 & 1 & 1 & 0 \\ 4 & 3 & 1 & 3 \end{pmatrix},$$

qui est équivalente par lignes, par les opérations $L_2 \leftarrow L_2 - 2L_1$, $L_3 \leftarrow L_3 - 3L_1$ et $L_4 \leftarrow L_4 - 4L_1$ à la matrice :

$$A_2 = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & -1 & -3 & -5 \\ 0 & -2 & -5 & -9 \\ 0 & -1 & -7 & -9 \end{pmatrix},$$

puis, par les opérations $L_3 \leftarrow L_3 - 2L_2$ et $L_4 \leftarrow L_4 - L_2$ à :

$$A_3 = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & -1 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -4 & -4 \end{pmatrix}$$

Cela permet de voir que la famille (a,b,c,d) est de rang 3, et que la famille (a,b,c) est libre. On a donc simultanément trouvé la dimension de F, qui est égale à 3, et trouvé une base de F, (a,b,c).

Nous avons déjà vu dans un exemple traité en **[SF2.13]** page 225 que trouver une base n'est pas toujours facile, et demande parfois un peu d'intuition. Pour cette raison, nous donnons un autre exemple un peu moins simple que le précédent.

Exemple

On reprend l'exemple de **[SF2.1]** où l'on a construit le complexifié $E \times E$ d'un $\mathbb{R}$-espace vectoriel E, et on suppose ici E de dimension finie.

Montrer que $\dim_{\mathbb{C}} E \times E = \dim_{\mathbb{R}} E$.

Soit n la dimension du $\mathbb{R}$-espace vectoriel E, et $(e_1,\ldots,e_n)$ une base de E. On sait qu'une base du $\mathbb{R}$-espace vectoriel produit $E \times E$ est formée des $2n$ couples $(e_i,0)$ et $(0,e_i)$, mais ici il faut trouver une base du $\mathbb{C}$-espace vectoriel $E \times E$. Or on peut remarquer que, si a et b sont des réels on a, par définition de la nouvelle loi externe :

$$(a + \mathrm{i}b).(e_k,0) = (ae_k,be_k) = a(e_k,0) + b(0,e_k)$$

ce qui suggère de montrer que la famille $\big((e_1,0),\ldots,(e_n,0)\big)$ est une base du $\mathbb{C}$-espace vectoriel $E \times E$.

- *Montrons que cette famille est génératrice de $E \times E$.*

  Pour tout $(x,y) \in E^2$ il existe des réels a_k et b_k tels que $x = \sum_{k=1}^{n} a_k e_k$ et $y = \sum_{k=1}^{n} b_k e_k$ d'où compte tenu du calcul précédent :

$$(x,y) = \sum_{k=0}^{n} \big(a_k(e_k,0) + b_k(0,e_k)\big) = \sum_{k=0}^{n} (a_k + \mathrm{i}b_k)(e_k,0)\,,$$

  ce qui démontre que la famille $\big((e_1,0),\ldots,(e_n,0)\big)$ est une famille génératrice du $\mathbb{C}$-espace vectoriel $E \times E$.

- *De plus cette famille est libre puisque le même calcul montre que :*

$$\sum_{k=0}^{n}(a_k+\mathrm{i}b_k)(e_k,0)=0_{E\times E}\implies\left(\sum_{k=1}^{n}a_ke_k,\sum_{k=1}^{n}b_ke_k\right)=(0,0)$$

ce qui implique que tous les a_k et tous les b_k, donc les $a_k+\mathrm{i}b_k$, sont nuls puisque $(e_1,\dots,e_n)$ est une famille libre du $\mathbb{R}$-espace vectoriel E.

- *En conclusion, la famille $\big((e_1,0),\dots,(e_n,0)\big)$ est une base du $\mathbb{C}$-espace vectoriel $E\times E$, qui est donc un $\mathbb{C}$-espace vectoriel de dimension $n=\dim_{\mathbb{R}}E$.*

SF2.19 Exhiber un isomorphisme avec un espace vectoriel connu

S'il existe un isomorphisme entre un espace vectoriel E et un espace vectoriel de dimension finie connu, alors les deux espaces ont la même dimension.

Exemple

Dans $\mathcal{M}_n(\mathbb{K})$, on considère l'ensemble E des matrices par blocs de la forme $\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}$ où A est une matrice carrée d'ordre r et B une matrice carrée d'ordre $n-r$ (avec r donné dans $[\![1\,;n-1]\!]$).

Montrer que E est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et en trouver la dimension.

- *On vérifie d'abord rapidement que E est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$.*
  - *E est bien inclus dans $\mathcal{M}_n(\mathbb{K})$, par construction.*
  - *E est non vide puisqu'il contient la matrice nulle.*
  - *Si $M=\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}$ et $M'=\begin{bmatrix}A' & 0\\ 0 & B'\end{bmatrix}$ (où A, A' sont carrées d'ordre r et B, B' carrées d'ordre $n-r$) alors les règles de calcul par blocs donnent, si λ est un scalaire : $\lambda M+M'=\begin{bmatrix}\lambda A+A' & 0\\ 0 & \lambda B+B'\end{bmatrix}$ donc $\lambda M+M'$ appartient à E.*

- *Considérons l'application* $\varphi:\left\{\begin{array}{ccc}E & \longrightarrow & \mathcal{M}_r(\mathbb{K})\times\mathcal{M}_{n-r}(\mathbb{K})\\ \begin{bmatrix}A & 0\\ 0 & B\end{bmatrix} & \longmapsto & (A,B)\,.\end{array}\right.$
  - *Il est facile de vérifier que φ est linéaire : les règles de calcul par blocs et la définition des lois dans un espace vectoriel produit donnent aisément $\varphi(\lambda M+M')=\lambda\varphi(M)+\varphi(M')$.*
  - *Soit $M=\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}$ dans E. Alors :*

$$\varphi(M)=0\iff A=0_r\text{ et }B=0_{n-r}\iff M=0_n\,,$$

ce qui montre que le noyau de φ est réduit à $\{0_n\}$; donc φ est injective.
  - *Enfin, φ est clairement surjective puisque tout couple $(A,B)\in\mathcal{M}_r(\mathbb{K})\times\mathcal{M}_{n-r}(\mathbb{K})$ est par construction de la forme $\varphi(M)$ pour $M\in E$.*

φ est donc un isomorphisme d'espaces vectoriels. Il en résulte que les espaces vectoriels de départ et d'arrivée ont la même dimension donc :

$$\begin{aligned}\dim E &= \dim(\mathcal{M}_r(\mathbb{K}) \times \mathcal{M}_{n-r}(\mathbb{K})) \\ &= \dim \mathcal{M}_r(\mathbb{K}) + \dim \mathcal{M}_{n-r}(\mathbb{K}) = r^2 + (n-r)^2 .\end{aligned}$$

✓ Il ne faut surtout pas oublier de vérifier que E est un espace vectoriel *avant* de parler d'application linéaire définie sur E : on ne peut parler d'isomorphisme qu'entre des espaces vectoriels !

De la même façon, il ne faut pas oublier de montrer qu'une application est linéaire *avant* de parler de son noyau !

SF2.20 Utiliser la formule de Grassmann

Exemple

Soient E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 2$ et H, H' deux hyperplans distincts de E.
Montrer que $H \cap H'$ est de dimension $n-2$.

- *Montrons d'abord que $E = H + H'$.*
  Puisque $H \neq H'$, il existe un vecteur a tel que $a \in H'$ et $a \notin H$. La droite vectorielle $\mathbb{K}.a$ est donc un supplémentaire de H d'après **[S2.20]**.
  On a ainsi $E = H \oplus \mathbb{K}.a$, et puisque $\mathbb{K}.a \subset H'$ on en déduit $E \subset H + H'$ puis l'égalité annoncée.
- *La formule de Grassmann donne alors :*

  $n = \dim E = \dim H + \dim H' - \dim(H \cap H') = 2(n-1) - \dim(H \cap H')$,

  d'où l'on tire immédiatement $\dim(H \cap H') = n-2$.

  ✓ *Ce résultat est un cas particulier d'un résultat plus général rappelé ci-dessous. Néanmoins il est d'usage assez fréquent, et sa démonstration est à connaître.*

SF2.21 Écrire un sous-espace comme intersection d'hyperplans

Cette méthode généralise la précédente. Il s'agit ici d'utiliser la caractérisation vue en **[S2.20]**, que nous rappelons de façon abrégée.

Dans un espace vectoriel de dimension n, l'intersection de p hyperplans « linéairement indépendants » est un sous-espace vectoriel de dimension $n-p$.

Exemple

Dans $\mathbb{R}^5$, on considère l'ensemble F des vecteurs (x_1,x_2,x_3,x_4,x_5) tels que :

$$\left\{\begin{array}{lcl} x_1 + 2x_2 - x_3 + 3x_4 + x_5 &=& 0 \\ x_2 + x_3 + 2x_4 - 2x_5 &=& 0 \\ x_2 + 2x_1 - 5x_3 - 4x_5 &=& 0 \end{array}\right. .$$

Montrer qu'il s'agit d'un sous-espace vectoriel de $\mathbb{R}^5$, en préciser la dimension et en trouver une base.

- *D'après* **[S2.20]**, *chacune des équations ci-dessus est celle d'un hyperplan ; F apparaît donc comme l'intersection de trois hyperplans, c'est donc un sous-espace vectoriel de $\mathbb{R}^5$.*

- *Pour déterminer la dimension de F, il suffit de déterminer le rang du système ci-dessus, afin de voir si les trois équations sont ou non indépendantes.*

  Le rang de ce système est celui de la matrice $\begin{pmatrix} 1 & 2 & -1 & 3 & 1 \\ 0 & 1 & 1 & 2 & -2 \\ 2 & 1 & -5 & 0 & -4 \end{pmatrix}$

  (***attention :*** *dans les équations données dans l'énoncé, les inconnues ne sont pas toutes dans le même ordre !).*

  L'opération $L_3 \longleftarrow L_3 - 2L_1 + 3L_2$ *montre que le rang de cette matrice est égal à celui de la matrice* $\begin{pmatrix} 1 & 2 & -1 & 3 & 1 \\ 0 & 1 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & -12 \end{pmatrix}$ *donc est égal à 3. Les 3 équations sont donc indépendantes, ce qui entraîne :* $\dim F = 5-3 = 2$.

- *Après l'opération* $L_3 \longleftarrow L_3 - 2L_1 + 3L_2$, *le système s'écrit :*

$$u = (x_1,\dots,x_5) \in F \iff \begin{cases} x_1 + 2x_2 - x_3 + 3x_4 + x_5 &= 0 \\ x_2 + x_3 + 2x_4 - 2x_5 &= 0 \\ -12x_5 &= 0 \end{cases}$$

$$\iff \begin{cases} x_5 &= 0 \\ x_2 &= -x_3 - 2x_4 \\ x_1 &= 3x_3 + x_4 \end{cases}$$

$$\iff u = (3x_3 + x_4, -x_3 - 2x_4, x_3, x_4, 0)$$
$$= x_3(3,-1,1,0,0) + x_4(1,-2,0,1,0)\,.$$

Cela montre que F est le sous-espace vectoriel de $\mathbb{R}^5$ engendré par les vecteurs $a = (3,-1,1,0,0)$ et $b = (1,-2,0,1,0)$. Ces deux vecteurs n'étant pas colinéaires, ils forment donc une base de F.

> ✓ *Pour trouver une base d'un sous-espace donné par un système d'équations, il suffit d'exprimer les inconnues secondaires en fonction des inconnues principales. Il faut retenir cette méthode, que nous avons utilisée ci-dessus.*

SF2.22 Utiliser le théorème du rang

Si l'on arrive à écrire un sous-espace vectoriel comme le noyau d'une application linéaire, le théorème du rang peut alors permettre d'obtenir sa dimension. La méthode précédente était en fait un cas particulier de cette méthode, puisque l'intersection d'hyperplans d'équations $\varphi_i(x) = 0$ est le noyau de l'application linéaire $x \mapsto (\varphi_1(x),\dots,\varphi_p(x))$ à valeurs dans $\mathbb{K}^p$.

Exemple

Soit E un espace vectoriel de dimension n, et $F_1,\dots,F_p$ des sous-espaces vectoriels de E tels que $\sum_{i=1}^{p} \dim F_i > n(p-1)$.

Montrer que $\bigcap_{i=1}^{p} F_i \neq \{0\}$.

$$\text{Soit } \psi : \begin{cases} F_1 \times \cdots \times F_p &\longrightarrow E^{p-1} \\ (x_1,\dots,x_p) &\longmapsto (x_2 - x_1, x_3 - x_2, \dots, x_p - x_{p-1})\,. \end{cases}$$

ψ est clairement une application linéaire.
Le noyau de ψ est formé des p-uplets $(x_1,\dots,x_p)$ tels que $x_1 = x_2 = \cdots = x_p$, c'est-à-dire des p-uplets de la forme $(x,x,\dots,x)$ avec $x \in \bigcap_{i=1}^{p} F_i$. Il est donc isomorphe à $\bigcap_{i=1}^{p} F_i$. Le théorème du rang implique alors :

$$\sum_{i=1}^{p} \dim F_i = \dim(F_1 \times \cdots \times F_p) = \operatorname{rg}\psi + \dim \bigcap_{i=1}^{p} F_i \,.$$

Puisque $\operatorname{rg}\psi \leqslant \dim E^{p-1} = n(p-1)$, l'hypothèse de l'énoncé implique $\dim \bigcap_{i=1}^{p} F_i > 0$, c'est-à-dire $\bigcap_{i=1}^{p} F_i \neq \{0\}$.

■ Comment démontrer l'égalité de deux sous-espaces vectoriels ?

SF2.23 Prouver une double inclusion

D'une façon générale, pour démontrer que deux ensembles A et B sont égaux, on peut démontrer successivement $A \subset B$ et $B \subset A$.

✓ Cette méthode est en fait assez peu utilisée : en effet, prouver une double inclusion est une méthode purement ensembliste, qui n'utilise pas toutes les propriétés des sous-espaces vectoriels.

Exemple

Soient A, B, C trois sous-espaces vectoriels d'un espace vectoriel E.

Montrer que : $A + \big(B \cap (A+C)\big) = A + \big(C \cap (A+B)\big)$.

Remarquons déjà qu'il suffit de prouver une inclusion, puisque la même démonstration en échangeant les rôles de B et C prouvera l'inclusion réciproque.
Montrons donc $A + \big(B \cap (A+C)\big) \subset A + \big(C \cap (A+B)\big)$.
Soit $x \in A + \big(B \cap (A+C)\big)$: x peut donc s'écrire $x = a + b$ avec $a \in A$ et $b \in B \cap (A+C)$, soit $b \in B$ et $b = a' + c$ avec $a' \in A$ et $c \in C$.
On a $c = (-a') + b$ donc $c \in A + B$ puis $c \in C \cap (A+B)$. Finalement $x = (a+a')+c$ avec $a+a' \in A$ et $c \in C \cap (A+B)$, donc $x \in A + \big(C \cap (A+B)\big)$, ce qui est le résultat voulu.

SF2.24 Prouver une inclusion puis l'égalité des dimensions

Lorsque deux sous-espaces vectoriels d'un même espace vectoriel sont de même dimension finie, il suffit de prouver une inclusion pour démontrer qu'ils sont égaux.

Exemple

Dans $\mathbb{R}^4$, soient $F = \big\{(x,y,z,t) \in \mathbb{R}^4 \;\big|\; x+y+z = 2x+y+z-t = 0\big\}$ et $G = \operatorname{Vect}\big((1,-2,1,1),(1,2,-3,1),(5,-3,-2,5)\big)$.

Montrer que $F = G$.

F est l'intersection des deux hyperplans distincts d'équations $x+y+z=0$ et $2x+y+z-t=0$, c'est donc un sous-espace vectoriel de $\mathbb{R}^4$ de dimension $4-2=2$.

Notons $a=(1,-2,1,1)$, $b=(1,2,-3,1)$ et $c=(5,-3,-2,5)$, de sorte que $G=\mathrm{Vect}(a,b,c)$. Puisque a et b ne sont évidemment pas colinéaires, G est de dimension $\geqslant 2$. Montrons que c est combinaison linéaire de a et b :

$$c=\alpha a+\beta b \iff \begin{cases} 5 &= \alpha+\beta \\ -3 &= -2\alpha+2\beta \\ -2 &= \alpha-3\beta \\ 5 &= \alpha+\beta \end{cases} \iff \begin{cases} 5 &= \alpha+\beta \\ -3 &= -2\alpha+2\beta \end{cases}$$

puisque la troisième ligne du premier système est équivalente à la somme des deux premières. On trouve donc $\alpha=\dfrac{13}{4}$ et $\beta=\dfrac{7}{4}$; c est donc bien combinaison linéaire de a et b, ce qui montre que $\dim G=2$.

Ainsi $\dim F=\dim G$. Pour démontrer l'égalité $F=G$ il suffit donc de montrer l'inclusion $G\subset F$. Pour cela il suffit de démontrer que les vecteurs a et b, qui forment une base de G, sont dans F, c'est-à-dire que leurs coordonnées vérifient les équations de F : cela est immédiat.

✓ *Il était ici bien plus facile de démontrer $G\subset F$ que $F\subset G$!*

■ Comment montrer que deux sous-espaces sont supplémentaires ?

SF2.25 Revenir à la définition

Il s'agit ici d'utiliser la définition :

$$E=F\oplus G \iff \forall\, x\in E,\ \exists!\,(y,z)\in F\times G \text{ tel que } x=y+z\,.$$

On cherche donc à écrire tout vecteur de E comme somme d'un vecteur de F et d'un vecteur de G, de façon unique ; pour cela, il est commode de procéder par *analyse-synthèse.*

Exemple

Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E tel que $u^3=\mathrm{Id}_E$.

Montrer que $E=\mathrm{Ker}(u-\mathrm{Id}_E)\oplus\mathrm{Im}(u-\mathrm{Id}_E)$.

- Analyse
  Soit $x\in E$; supposons qu'il existe x_1, x_2 avec $x_1\in\mathrm{Ker}(u-\mathrm{Id}_E)$ et $x_2\in\mathrm{Im}(u-\mathrm{Id}_E)$ tels que $x=x_1+x_2$. Par définition, on a $u(x_1)=x_1$ et il existe $y\in E$ tel que $x_2=u(y)-y$.
  On a donc $x=x_1+u(y)-y$ d'où $u(x)=x_1+u^2(y)-u(y)$ puis $u^2(x)=x_1+u^3(y)-u^2(y)=x_1+y-u^2(y)$.
  En additionnant ces 3 relations, on trouve $x_1=\dfrac{1}{3}\big(u^2(x)+u(x)+x\big)$. Si x_1 existe, il est donc unique, et par suite $x_2=x-x_1$ aussi.
- Synthèse
  Réciproquement, soit $x\in E$. Posons $x_1=\dfrac{1}{3}\big(u^2(x)+u(x)+x\big)$, puis $x_2=x-x_1$.

– *On a* $u(x_1) = \frac{1}{3}\big(u^3(x) + u^2(x) + u(x)\big) = x_1$, *donc* $x_1 \in \mathrm{Ker}(u - \mathrm{Id}_E)$.

– *On remarque alors que :*

$$3x_1 = u^2(x) + u(x) + x = (u - \mathrm{Id}_E)\big(u(x) + 2x\big) + 3x$$

donc $x_2 = x - x_1 = (u - \mathrm{Id}_E)\left(\frac{-u(x) - 2x}{3}\right) \in \mathrm{Im}(u - \mathrm{Id}_E)$.

- *En conclusion : tout* $x \in E$ *s'écrit de façon unique comme somme d'un vecteur de* $\mathrm{Ker}(u - \mathrm{Id}_E)$ *et d'un vecteur de* $\mathrm{Im}(u - \mathrm{Id}_E)$, *ce qui prouve que ces deux sous-espaces vectoriels sont supplémentaires.*

  ✓ *Lors de la synthèse, nous avons en fait seulement prouvé l'inclusion* $E \subset \mathrm{Ker}(u - \mathrm{Id}_E) + \mathrm{Im}(u - \mathrm{Id}_E)$. *Mais l'inclusion réciproque est évidente, et pour cette raison, on ne la mentionne jamais.*

SF2.26 Étudier l'intersection et la somme

Il s'agit ici d'utiliser la caractérisation :

$$E = F \oplus G \iff E = F + G \text{ et } F \cap G = \{0\}.$$

Cette méthode n'est pas forcément plus simple que la précédente : en effet, lorsque l'on recherche la décomposition d'un vecteur de E comme somme d'un vecteur de F et d'un vecteur de G en procédant par *analyse*, on obtient en général directement l'unicité de cette décomposition (voir l'exemple précédent).
Il faut donc plutôt utiliser cette méthode lorsque l'on connaît déjà la décomposition ; alors, au lieu de montrer son unicité, on peut simplement vérifier $F \cap G = \{0\}$.

Exemple

Dans l'espace vectoriel E des fonctions continues sur $[0\,;1]$ à valeurs dans $\mathbb{R}$, soit :

$$H = \left\{ f \in E \;\middle|\; \int_0^1 f(t)\,\mathrm{d}t = 0 \right\}.$$

Trouver un supplémentaire de H dans E.

- *Là encore, même si ce n'est pas dit dans l'énoncé, il ne faut pas oublier de vérifier que* H *est un sous-espace vectoriel de* E.
  Pour ce faire, on peut, soit revenir à la caractérisation d'un sous-espace vectoriel, soit plutôt remarquer que H *est le noyau de la forme linéaire* $f \mapsto \int_0^1 f$ *; le résultat est alors immédiat.*
- *Si* $f \in E$, *la fonction* $f - \int_0^1 f$ *est d'intégrale nulle sur* $[0\,;1]$, *donc appartient à* H. *On peut donc écrire toute fonction* f *de* E *sous la forme* $f = \left(f - \int_0^1 f\right) + \int_0^1 f$ *c'est-à-dire comme somme d'une fonction de* H *et d'une fonction constante. Ainsi, en notant* D *l'espace vectoriel des applications constantes sur* $[0\,;1]$, *on a* $E = H + D$.
- *Il est alors facile de vérifier que* $H \cap D = \{0\}$ *(une fonction constante d'intégrale nulle est nulle) donc* $E = H \oplus D$.

SF2.27 Étudier l'intersection et les dimensions

Dans un espace vectoriel de dimension finie, on peut utiliser la caractérisation :

$$E = F \oplus G \iff F \cap G = \{0\} \quad \text{et} \quad \dim E = \dim F + \dim G\,.$$

Il s'agit certainement de la méthode à privilégier dans un espace vectoriel de dimension finie. Son principal défaut est de n'être pas constructive, c'est-à-dire de ne pas expliciter la décomposition d'un vecteur de E comme somme de vecteurs de F et de G, mais, justement, il y a des cas où cela n'est pas utile.

Exemple

Dans $\mathbb{R}^4$, soient $F = \{(x,y,z,t) \mid x+y+z+t = x-y+z-t = 0\}$
et $G = \{(x,y,z,t) \mid x+2y+3z+4t = x-2y+3z-4t = 0\}$.

Démontrer que F et G sont deux sous-espaces vectoriels supplémentaires de $\mathbb{R}^4$.

Tout d'abord, comme il a déjà été dit plusieurs fois, il est facile de montrer que F et G sont des plans vectoriels : en effet, chacun d'entre eux est l'intersection de deux hyperplans distincts de $\mathbb{R}^4$.

On a donc bien $\dim \mathbb{R}^4 = \dim F + \dim G$. Pour répondre à la question, il suffit donc de montrer que $F \cap G = \{0\}$.
Soit donc $u \in F \cap G$. Les coordonnées (x,y,z,t) de u vérifient le système :

$$\begin{cases} x+y+z+t &= 0 \\ x-y+z-t &= 0 \\ x+2y+3z+4t &= 0 \\ x-2y+3z-4t &= 0 \end{cases}.$$

En additionnant et en soustrayant les deux premières lignes, et de même pour les deux dernières, on obtient :

$$x+z=0\ ;\ y+t=0\ ;\ x+3z=0\ ;\ y+2t=0$$

d'où $x=y=z=t=0$ soit $u=0$ puis le résultat cherché.

■ Comment montrer que n $(n \geqslant 3)$ sous-espaces sont supplémentaires ?

Les principales définitions et propriétés sont rappelées en **[S2.10]** et en **[S2.14]**.
Pour montrer que la somme de sous-espaces $E_1, \ldots, E_n$ est directe, on ne peut pas se contenter de démontrer, comme dans le cas $n=2$, que $E_i \cap E_j = \{0\}$ pour $i \neq j$.

✓ Attention à ne pas commettre cette erreur ! Ayez toujours présent à l'esprit ce contre-exemple simple : on considère trois droites vectorielles distinctes dans le plan. Leurs intersections 2 à 2 sont réduites à $\{0\}$, mais la somme des 3 droites ne peut être directe pour des raisons de dimension.

SF2.28 Revenir à la définition

On utilise la caractérisation suivante.

Les sous-espaces $(E_i)_{1 \leqslant i \leqslant n}$ sont supplémentaires dans E si et seulement si : tout vecteur de E s'écrit de façon unique comme somme de vecteurs des E_i.

On peut bien sûr procéder en deux temps : d'abord montrer l'existence de la décomposition, et ensuite l'unicité de celle-ci, qui équivaut à :

$$\forall\,(x_1,\dots,x_n)\in E_1\times\dots\times E_n,\ x_1+\dots+x_n=0\implies x_1=\dots=x_n=0\,.$$

Exemple

Soit E un $\mathbb{C}$-espace vectoriel, et $f\in\mathscr{L}(E)$ tel que $f^3=\mathrm{Id}_E$.

On note j le nombre complexe $\mathrm{e}^{\frac{2\mathrm{i}\pi}{3}}$ et on pose :

$$f_1=\mathrm{Id}_E+f+f^2\quad,\quad f_2=\mathrm{Id}_E+\mathrm{j}f+\mathrm{j}^2f^2\quad\text{et}\quad f_3=\mathrm{Id}_E+\mathrm{j}^2f+\mathrm{j}f^2\,.$$

Montrer que f est un automorphisme de E et que $E=\bigoplus_{k=1}^{3}\mathrm{Im}\,f_k$.

- *$f^2\circ f=f\circ f^2=\mathrm{Id}_E$ donc f est bijectif et $f^{-1}=f^2$.*
- *Puisque $1+\mathrm{j}+\mathrm{j}^2=0$, tout x de E s'écrit : $x=\frac{1}{3}\big(f_1(x)+f_2(x)+f_3(x)\big)$, d'où $E=\sum_{k=1}^{3}\mathrm{Im}\,f_k$.*
- *Pour $i\in[\![1\,;3]\!]$, soit $x_i\in\mathrm{Im}\,f_i$ tels que $x_1+x_2+x_3=0$. Pour tout i, il existe $y_i\in E$ tel que $x_i=f_i(y_i)$. On a ainsi :*

  $$f_1(y_1)+f_2(y_2)+f_3(y_3)=0\quad(*)$$

  On remarque alors que, puisque $f^3=\mathrm{Id}_E$ et $\mathrm{j}^3=1$, $f\circ f_1=f_1$, $f\circ f_2=\mathrm{j}^2f_2$ et $f\circ f_3=\mathrm{j}f_3$.
  En appliquant alors f puis f^2 à l'égalité $(*)$ on obtient le système :

  $$\begin{cases}f_1(y_1)+f_2(y_2)+f_3(y_3)&=0\\ f_1(y_1)+\mathrm{j}^2f_2(y_2)+\mathrm{j}f_3(y_3)&=0\\ f_1(y_1)+\mathrm{j}f_2(y_2)+\mathrm{j}^2f_3(y_3)&=0\end{cases}$$

  qui donne facilement $f_1(y_1)=f_2(y_2)=f_3(y_3)=0$ (la somme des trois équations donne $f_1(y_1)=0$, puis l'opération $L_1+\mathrm{j}L_2+\mathrm{j}^2L_3$ donne $f_2(y_2)=0$) soit $x_i=0$ pour tout i, et la somme est directe.

SF2.29 Utiliser les dimensions

Lorsque E est de dimension finie, on peut utiliser la caractérisation suivante.

Les sous-espaces $(E_i)_{1\leqslant i\leqslant n}$ sont supplémentaires dans E si et seulement si :

- $\forall\,(x_1,\dots,x_n)\in E_1\times\dots\times E_n,\ x_1+\dots+x_n=0\implies x_1=\dots=x_n=0\,,$
- et $\dim E=\sum_{i=1}^{n}\dim E_i\,.$

Exemple

Dans $E=\mathbb{R}_3[X]$ on considère les sous-ensembles :

$$F=\{P\in E\mid P(0)=P(1)=P(2)=0\}\,,$$
$$G=\{P\in E\mid P(1)=P(2)=P(3)=0\}\quad\text{et}\quad H=\{P\in E\mid P(X)=P(-X)\}\,.$$

Montrer qu'il s'agit de trois sous-espaces vectoriels supplémentaires de E.

Bien sûr, il ne faut pas oublier de montrer qu'il s'agit bien de sous-espaces vectoriels de E, ce que nous ferons dans le cours de la démonstration.

- *F est l'ensemble des polynômes qui s'annulent en 0, 1 et 2, donc qui sont multiples de $X(X-1)(X-2)$. Puisque les polynômes considérés ici sont de degré $\leqslant 3$, il s'agit des polynômes de la forme $\lambda X(X-1)(X-2)$ lorsque λ décrit $\mathbb{R}$, donc de la droite vectorielle engendrée par $X(X-1)(X-2)$.*
- *De même, G est la droite vectorielle engendrée par $(X-1)(X-2)(X-3)$.*
- *H est l'ensemble des polynômes de degré $\leqslant 3$ et pairs, donc de la forme $a+bX^2$: il s'agit du sous-espace vectoriel* $\text{Vect}\left(1,X^2\right)$.
- *Montrons que la somme $F+G+H$ est directe. Soit $(P,Q,R) \in F\times G\times H$ avec $P+Q+R=0$. Il existe des scalaires λ, μ, a et b tels que :*

  $$P=\lambda X(X-1)(X-2)\,,\ Q=\mu(X-1)(X-2)(X-3)\,,\ R=a+bX^2\,,$$

  donc la relation $P+Q+R=0$ s'écrit :

  $$\lambda X(X-1)(X-2)+\mu(X-1)(X-2)(X-3)+a+bX^2=0\,,$$

  ce qui conduit facilement, en appliquant successivement cette relation en 0, 1, 2 puis 3, à $\lambda=\mu=a=b=0$ donc à $P=Q=R=0$.
- *Enfin, l'égalité $\dim F+\dim G+\dim H = 1+1+2 = 4 = \dim \mathbb{R}_3[X]$ permet de conclure.*

> ✓ *Nous aurions pu aussi démontrer le résultat en utilisant l'associativité de la somme directe, c'est-à-dire en démontrant déjà $F\cap G=\{0\}$, donc la somme $F+G$ est directe, puis que $(F\oplus G)\cap H=\{0\}$. En fait seule la présentation diffère, les calculs sont similaires à ceux ci-dessus.*

SF2.30 Utiliser une base adaptée

En dimension finie, on peut aussi démontrer que la réunion de bases des E_i forme une base de E.

Exemple

Reprenons l'exemple précédent.

On avait trouvé des bases des sous-espaces vectoriels F, G et H. Il suffit donc de montrer que la réunion de ces bases, c'est-à-dire la famille $\left(X(X-1)(X-2),(X-1)(X-2)(X-3),1,X^2\right)$ est une base de $\mathbb{R}_3[X]$.

Pour cela, on peut établir qu'il s'agit d'une famille libre, ce qui revient à faire le même genre de calculs que dans l'exemple précédent, ou bien démontrer que le déterminant de ces quatre vecteurs dans la base canonique $\left(1,X,X^2,X^3\right)$ de $\mathbb{R}_3[X]$ est non nul. Or puisque :

$$X(X-1)(X-2)=X^3-3X^2+2X$$
$$\text{et}\quad (X-1)(X-2)(X-3)=X^3-6X^2+11X-6\,,$$

ce déterminant est égal à : $\Delta=\begin{vmatrix} 0 & -6 & 1 & 0\\ 2 & 11 & 0 & 0\\ -3 & -6 & 0 & 1\\ 1 & 1 & 0 & 0\end{vmatrix}=9$, *d'où le résultat.*

SF2.31 Utiliser le théorème de décomposition des noyaux

Lorsque l'énoncé fait apparaître un polynôme d'endomorphisme, ce théorème rappelé en **[S4.9]** peut être très utile. On pouvait d'ailleurs l'utiliser pour simplifier les démonstrations des exemples pages 234 et 237. Nous ne donnons pas d'autre exemple pour l'instant, nous le reverrons lors du chapitre concernant la réduction.

■ Comment montrer que u est une application linéaire de E dans F ?

SF2.32 Utiliser la définition ou les propriétés

Pour montrer que u est une application linéaire de E dans F, la seule difficulté, parfois cachée, est de démontrer que u est bien une *application* de E dans F, c'est-à-dire que u est bien définie sur E entier et que son image est contenue dans F.

La linéarité résulte alors d'un calcul souvent immédiat, mais on peut parfois aussi utiliser le fait que la somme ou la composée d'applications linéaires est linéaire.

Exemple

Soit $E_0 = \{f \in \mathscr{C}(\mathbb{R},\mathbb{R}) \mid f(0) = 0\}$. Pour toute $f \in E_0$, on considère l'application $T(f)$ définie par : $\forall\, x \in \mathbb{R}^*,\ T(f)(x) = \dfrac{1}{x}\displaystyle\int_0^x tf(t)\,\mathrm{d}t$ et $T(f)(0) = 0$.

Montrer que E_0 est un $\mathbb{R}$-espace vectoriel et que T est un endomorphisme de E_0.

- *E_0 est le noyau de la forme linéaire $f \mapsto f(0)$, c'est donc un hyperplan de $\mathscr{C}(\mathbb{R},\mathbb{R})$, et en particulier un sous-espace vectoriel de $\mathscr{C}(\mathbb{R},\mathbb{R})$.*
- *Soit $f \in E_0$. f est continue sur $\mathbb{R}$ donc la fonction $G : x \mapsto \displaystyle\int_0^x tf(t)\,\mathrm{d}t$ est une primitive de la fonction $t \mapsto tf(t)$; G est donc en particulier continue sur $\mathbb{R}$, et $T(f)$ sera continue sur $\mathbb{R}^*$.*

  De plus, pour $x \neq 0$, $T(f)(x) = \dfrac{G(x)-G(0)}{x} \underset{x\to 0}{\longrightarrow} G'(0)$ et $G'(0) = 0$ puisque $G'(x) = xf(x)$. Ainsi $\lim\limits_{x\to 0} T(f)(x) = T(f)(0)$: $T(f)$ est continue en 0, et par suite sur $\mathbb{R}$.
  Enfin, puisque $T(f)(0) = 0$ on a bien $T(f) \in E_0$.
- *La linéarité de T découle immédiatement de celle de l'intégration.*

Finalement, T est bien un endomorphisme de E_0.

✓ *Le plus difficile dans cet exemple n'était pas de montrer la linéarité de T, mais de vérifier que T est bien une application de E_0 dans E_0 !*

■ Comment construire une application linéaire ?

SF2.33 Donner l'image de tous les vecteurs

C'est évidemment la méthode de base : on se donne une expression permettant de calculer $u(x)$ pour tout $x \in E$. Bien sûr, il ne faut pas oublier de vérifier que

cette définition a bien un sens et donne bien une application linéaire ; ce n'est pas toujours si simple, comme le montrent les exemples suivants.
Lorsque E est de dimension finie, muni d'une certaine base, cela revient à se donner l'expression analytique de u, ou encore sa matrice.

Exemple

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels, $v \in \mathscr{L}(F,G)$ et $w \in \mathscr{L}(E,G)$ telles que $\operatorname{Im} w \subset \operatorname{Im} v$.
Démontrer qu'il existe une application linéaire $u \in \mathscr{L}(E,F)$ telle que $w = v \circ u$.

Soit S un supplémentaire de $\operatorname{Ker} v$ dans F. D'après le théorème d'isomorphisme rappelé en **[S2.17]**, *la restriction de v à S induit un isomorphisme v_1 entre S et $\operatorname{Im} v$.*
L'application $v_1^{-1} \circ w$ est bien définie puisque $\operatorname{Im} w \subset \operatorname{Im} v$, et elle est bien linéaire comme composée de deux applications linéaires. Elle est définie de E dans S, mais on peut bien sûr la considérer comme une application linéaire de E dans F ; notons là u.
On a alors, pour tout $x \in E$:

$$v \circ u(x) = v[u(x)] \underbrace{=}_{\substack{car \\ u(x) \in S}} v_1[u(x)] = v_1 \circ v_1^{-1} \circ w(x) = w(x)\,,$$

donc u répond bien à la question.

Exemple

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels, $u \in \mathscr{L}(E,F)$ et $w \in \mathscr{L}(E,G)$ telles que $\operatorname{Ker} u \subset \operatorname{Ker} w$.
Démontrer qu'il existe une application linéaire $v \in \mathscr{L}(F,G)$ telle que $w = v \circ u$.

- *Soit $y \in \operatorname{Im} u$: il existe $x \in E$ tel que $y = u(x)$; posons alors $v_1(y) = w(x)$.*
  - *Cette définition a bien un sens : en effet, si x et x' sont deux antécédents de y par u, c'est-à-dire $y = u(x) = u(x')$, alors $u(x - x') = 0$ donc $x - x' \in \operatorname{Ker} u$ donc $x - x' \in \operatorname{Ker} w$ par hypothèse. On aura donc aussi $w(x) = w(x')$, et la définition de $v_1(y)$ est cohérente.*
  - *Montrons que v_1 est une application linéaire de $\operatorname{Im} u$ dans G : soient y, $y' \in \operatorname{Im} u$ et $\lambda \in \mathbb{K}$. Il existe x, $x' \in E$ tels que $y = u(x)$ et $y' = u(x')$ donc $\lambda y + y' = u(\lambda x + x')$ puisque u est linéaire. Par définition on aura donc :*

    $$v_1(\lambda y + y') = w(\lambda x + x') \underbrace{=}_{\substack{w \\ linéaire}} \lambda w(x) + w(x') = \lambda v_1(y) + v_1(y')\,,$$

    ce qui prouve la linéarité de v_1.
- *Par construction on a : $\forall\, x \in E$, $w(x) = v_1 \circ u(x)$. Enfin, n'importe quelle application linéaire v de F dans G dont la restriction à $\operatorname{Im} u$ est égale à v_1 (il en existe, puisqu'il suffit de la définir sur un supplémentaire de $\operatorname{Im} u$) vérifie $w = v \circ u = v_1 \circ u$ donc convient.*

SF2.34 La définir sur des sous-espaces vectoriels supplémentaires

Il s'agit ici d'utiliser **[S2.18]** : une application linéaire est entièrement déterminée si l'on connaît ses restrictions à des sous-espaces vectoriels supplémentaires.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathscr{L}(E)$.
Soit A l'ensemble des endomorphismes v de E tels que $u \circ v = v \circ u = 0$.

Montrer que A est un sous-espace vectoriel de $\mathscr{L}(E)$ et en donner la dimension.

- *Le fait que A est un sous-espace vectoriel de $\mathscr{L}(E)$ est facile.*
  - *A est bien inclus dans $\mathscr{L}(E)$ par définition.*
  - *A est non vide car il contient évidemment l'application nulle.*
  - *Si v_1, v_2 appartiennent à A et si λ est un scalaire alors*
$$(\lambda v_1 + v_2) \circ u = \lambda v_1 \circ u + v_2 \circ u = 0$$
$$u \circ (\lambda v_1 + v_2) = \lambda u \circ v_1 + u \circ v_2 = 0\,,$$
  donc $\lambda v_1 + v_2$ appartient à A.
- *Notons déjà que :*
$$v \in A \iff \operatorname{Im} v \subset \operatorname{Ker} u \quad \text{et} \quad \operatorname{Im} u \subset \operatorname{Ker} v\,.$$
Considérons un supplémentaire S de $\operatorname{Im} u$ dans E. À toute application v appartenant à A on peut associer l'application $v_1 \in \mathscr{L}(S, \operatorname{Ker} u)$ telle que, pour tout $x \in S$, $v_1(x) = v(x)$.
L'application $\varphi : \begin{cases} A & \longrightarrow & \mathscr{L}(S, \operatorname{Ker} u) \\ v & \longmapsto & v_1 \end{cases}$ est clairement linéaire.
Réciproquement, soit $v_1 \in \mathscr{L}(S, \operatorname{Ker} u)$; en vertu de ce qui a été rappelé ci-dessus, on peut alors construire une application linéaire v de E dans E en posant :
$$\forall\, x \in S,\ v(x) = v_1(x) \quad \text{et} \quad \forall\, x \in \operatorname{Im} u,\ v(x) = 0\,.$$
On aura par construction $v \in A$ et $\varphi(v) = v_1$. Ainsi, l'application φ est surjective.
Elle est également injective : si $\varphi(v) = 0$ alors les restrictions de v aux sous-espaces vectoriels supplémentaires S et $\operatorname{Im} u$ sont nulles donc $v = 0$.
φ est donc un isomorphisme entre A et $\mathscr{L}(S, \operatorname{Ker} u)$; il en résulte que :
$$\dim A = \dim\big(\mathscr{L}(S, \operatorname{Ker} u)\big) = \dim S \times \dim \operatorname{Ker} u = (\dim \operatorname{Ker} u)^2$$
(en effet, $\dim S = \dim \operatorname{Ker} u$ en vertu du théorème du rang).

SF2.35 Donner l'image d'une base

Il s'agit ici d'utiliser une autre caractérisation rappelée en **[S2.18]** : une application linéaire sur un espace vectoriel de dimension finie est entièrement déterminée si l'on connaît l'image d'une base.

Exemple

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, et F, G deux sous-espaces vectoriels de E tels que $\dim F + \dim G = n$.

Montrer qu'il existe $u \in \mathscr{L}(E)$ tel que $\operatorname{Ker} u = F$ et $\operatorname{Im} u = G$.

- *Si F et G sont supplémentaires, c'est-à-dire $F \cap G = \{0\}$, la projection sur G parallèlement à F convient.*
- *Sinon, soient $r = \dim(F \cap G)$, $p = \dim F$ et $q = \dim G$.*
  Soit $(e_1,\dots,e_r)$ une base de $F \cap G$. On peut la compléter en une base $(e_1,\dots,e_p)$ de F en choisissant $(e_{r+1},\dots,e_p)$ base d'un supplémentaire F' de $F \cap G$ dans F, et en une base $(e_1,\dots,e_r,e'_{r+1},\dots,e'_q)$ de G en choisissant $(e'_{r+1},\dots,e'_q)$ base d'un supplémentaire G' de $F \cap G$ dans G.
  Puisque $F+G = (F \cap G) \oplus F' \oplus G'$, la famille $(e_1,\dots,e_p,e'_{r+1},\dots,e'_q)$ est une base de $F+G$; elle possède $p+q-r$ éléments donc, puisque $n = p+q$ d'après l'hypothèse de l'énoncé, il existe $(e''_1,\dots,e''_r)$ tels que :
  $$(e_1,\dots,e_p,e'_{r+1},\dots,e'_q,e''_1,\dots,e''_r)$$
  soit une base de E.
  On sait alors qu'il existe un endomorphisme u de E qui transforme cette base en la famille :
  $$(0,\dots,0,e'_{r+1},\dots,e'_q,e_1,\dots,e_r)$$
  et l'on aura bien, par construction : $\operatorname{Ker} u = F$ *et* $\operatorname{Im} u = G$.

■ Comment déterminer l'image ou le rang d'une application linéaire ?

SF2.36 Revenir à la définition

Il s'agit ici d'utiliser la définition purement ensembliste de l'image, à savoir :

$$\operatorname{Im} u = \{u(x),\ x \in E\}\,.$$

Cette méthode, qui n'utilise pas en général les propriétés liées à la linéarité, est plutôt à éviter, sauf parfois en dimension infinie.

Exemple

Déterminer l'image de l'application u qui à tout $f \in \mathscr{C}(\mathbb{R},\mathbb{R})$ associe la fonction :

$$u(f) : x \in \mathbb{R} \longmapsto \int_0^x f(t)\,\mathrm{d}t\,.$$

- *Notons d'abord que la définition de u a bien un sens (f étant continue, l'intégrale existe), et que u est linéaire d'après les propriétés de l'intégrale.*
- *On remarque alors que pour tout $f \in \mathscr{C}(\mathbb{R},\mathbb{R})$, $u(f)$ est une application de classe $\mathscr{C}^1$, puisque $\big(u(f)\big)' = f$, et que $u(f)$ s'annule en 0.*
  Donc l'image de u est incluse dans l'ensemble :
  $$E = \left\{g \in \mathscr{C}^1(\mathbb{R},\mathbb{R}) \;\middle|\; g(0) = 0\right\}\,.$$
- *Réciproquement, si g appartient à E, on a :*
  $$\forall\, x \in \mathbb{R},\ \int_0^x g'(t)\,\mathrm{d}t = g(x) - g(0) = g(x)$$
  avec g' continue, donc $g = u(g')$ et g appartient à $\operatorname{Im} u$.

En conclusion, l'image de u est exactement égale à E (et cela prouve que E est un espace vectoriel, ce que nous n'avions pas démontré).

SF2.37 Utiliser une base

On rappelle que l'image d'une application linéaire est aussi le sous-espace vectoriel engendré par les images d'une famille génératrice (ou d'une base). Cette méthode permet d'ailleurs de déterminer, non seulement l'image de l'espace tout entier, mais aussi celle de n'importe quel sous-espace vectoriel.

En anticipant sur le chapitre suivant, c'est d'ailleurs cette propriété qui est utilisée lorsque l'on recherche l'image d'une application linéaire dont on connaît la matrice dans une base, puisque les images des vecteurs de base sont justement données par les colonnes de la matrice.

Exemple

Déterminer l'image de $u : \begin{cases} \mathbb{R}[X] & \longrightarrow & \mathbb{R}[X] \\ P & \longmapsto & P(X+1) - P(X) \end{cases}$.

- *Il est facile de vérifier que u est bien un endomorphisme de l'espace vectoriel $\mathbb{R}[X]$.*
  - *u est bien une application de $\mathbb{R}[X]$ dans $\mathbb{R}[X]$.*
  - *Pour tous polynômes P et Q et tout scalaire λ, on a facilement $u(\lambda P + Q) = \lambda u(P) + u(Q)$ (calcul immédiat), donc u est linéaire.*
- *Une base (infinie) de $\mathbb{R}[X]$ est la famille $\left(X^n\right)_{n\in\mathbb{N}}$ (base canonique). L'image de u est donc le sous-espace vectoriel de $\mathbb{R}[X]$ engendré par les polynômes $u(X^n)$ pour $n \in \mathbb{N}$.*
  Or pour $n = 0$, $u(1) = 0$, et pour $n \geqslant 1$:
  $$u(X^n) = (X+1)^n - X^n = \sum_{k=0}^{n} \binom{n}{k} X^k - X^n = \sum_{k=0}^{n-1} \binom{n}{k} X^k$$
  est un polynôme de degré $n-1$.
  En utilisant le résultat du premier exemple de **[SF2.13]**, *on obtient que la famille des polynômes $u(X^n)$ pour $n \in \mathbb{N}^*$ est une base de $\mathbb{K}[X]$, donc la famille $\left(u(X^n)\right)_{n\in\mathbb{N}}$ engendre $\mathbb{K}[X]$: cela prouve que l'image de u est égale à $\mathbb{K}[X]$ (u est surjective).*

SF2.38 Utiliser le théorème du rang

Dans l'immense majorité des cas, il est plus facile de déterminer le noyau que l'image d'une application linéaire, puisqu'il s'agit alors de résoudre une « simple » équation $u(x) = 0$. Lorsque l'espace vectoriel est de dimension finie, le théorème du rang (rappelé en **[S2.19]**) permet d'avoir un renseignement concernant l'image si l'on a préalablement déterminé le noyau.

Quelques remarques à propos de ce théorème, qui est certainement l'un des plus populaires en algèbre linéaire, mais aussi l'un des plus mal utilisés.

- Le théorème du rang n'est qu'une application, en dimension finie, d'un théorème plus général et plus précis, le théorème d'isomorphisme (voir **[S2.17]**).
- Dans les hypothèses du théorème du rang, seul l'espace de départ a besoin d'être de dimension finie. À force d'appliquer ce théorème dans le cas d'endomorphismes, beaucoup ne savent d'ailleurs plus si la dimension de l'espace vectoriel qui intervient est celle de l'espace de départ ou d'arrivée !

– Dans le cas d'un endomorphisme d'un espace vectoriel de dimension finie, le théorème du rang ne dit **pas** que image et noyau sont supplémentaires : il n'y a aucune raison que leur intersection soit réduite à $\{0\}$.

Exemple

Soit n un entier $\geqslant 2$ et u l'application qui à tout polynôme P de $\mathbb{K}_n[X]$ associe le polynôme $u(P) = XP(1) + X^2P(0)$.

Montrer que u est un endomorphisme de $\mathbb{K}_n[X]$. Déterminer son image et son noyau.

- *Notons déjà que la définition a bien un sens puisque, si P est un polynôme de $\mathbb{K}_n[X]$, $u(P)$ est encore un polynôme, et $u(P) \in \mathbb{K}_n[X]$ puisque $n \geqslant 2$.*
  Montrons que u est linéaire : pour tous P, $Q \in \mathbb{R}_n[X]$ et tout $\lambda \in \mathbb{R}$:
  $$\begin{aligned} u(\lambda P + Q) &= X(\lambda P + Q)(1) + X^2(\lambda P + Q)(0) \\ &= X(\lambda P(1) + Q(1)) + X^2(\lambda P(0) + Q(0)) \\ &= \lambda(XP(1) + X^2P(0)) + (XQ(1) + X^2Q(0)) \\ &= \lambda u(P) + u(Q)\,. \end{aligned}$$
  ✓ *Nous avons encore fait la démonstration complète de la linéarité de u, mais cela ne devrait plus vous poser de problème !*
- $u(P) = 0 \iff XP(1) + X^2P(0) = 0 \iff P(1) = P(0) = 0$. *Donc le noyau de u est formée des polynômes de $\mathbb{K}_n[X]$ admettant 0 et 1 pour racines, c'est-à-dire divisibles par $X(X-1)$:*
  $$P \in \operatorname{Ker} u \iff \exists\, Q \in \mathbb{K}_{n-2}[X] \text{ tel que } P = X(X-1)Q\,.$$
  $\operatorname{Ker} u$ *est donc isomorphe à $\mathbb{K}_{n-2}[X]$, et est de dimension $n-1$.*
- *Le théorème du rang permet alors de dire que le rang de u est égal à* $\dim \mathbb{K}_n[X] - \dim \operatorname{Ker} u = 2$.
  Puisque de façon évidente $\operatorname{Im} u \subset \operatorname{Vect}(X, X^2)$, *et que* $\dim \operatorname{Im} u = \dim \operatorname{Vect}(X, X^2) = 2$, *on aura* $\operatorname{Im} u = \operatorname{Vect}(X, X^2)$.

SF2.39 Utiliser des inégalités sur le rang

On peut parfois obtenir des renseignements en utilisant les relations sur le rang rappelées en **[S2.19]**.

Exemple

Soit p un projecteur de $\mathbb{R}^3$, de rang 2. On suppose qu'il existe $u \in \mathscr{L}(\mathbb{R}^2, \mathbb{R}^3)$ et $v \in \mathscr{L}(\mathbb{R}^3, \mathbb{R}^2)$ tels que $u \circ v = p$.

Que peut-on dire de $v \circ u$?

Puisque $\operatorname{rg}(u \circ v) \leqslant \operatorname{rg} u$, *on a* $\operatorname{rg} u \geqslant \operatorname{rg} p$ *d'où* $\operatorname{rg} u \geqslant 2$. *Mais puisque u est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^3$, on a aussi* $\operatorname{rg} u \leqslant \dim \mathbb{R}^2 = 2$ *et finalement* $\operatorname{rg} u = 2$.
Mais alors, puisque $\operatorname{rg} u = \dim \mathbb{R}^2$ *(dimension de l'espace de départ), on en déduit que u est injective.*

On démontre de la même manière que v est surjective.

De plus, p est un projecteur (vous remarquerez que l'on ne s'est pas encore servi de cette hypothèse !) donc $p \circ p = p$ soit $u \circ v \circ u \circ v = u \circ v$.
On a donc $u\big[v \circ u \circ v(x)\big](x) = u\big[v(x)\big]$ pour tout $x \in E$; u étant injective, cela implique $v \circ u \circ v(x) = v(x)$ pour tout $x \in E$.
v étant surjective, cela implique, pour tout $y \in E$: $v \circ u(y) = y$.
Finalement, on peut conclure : $v \circ u = \mathrm{Id}_{\mathbb{R}^2}$.

■ Comment déterminer le noyau d'une application linéaire ?

SF2.40 Revenir à la définition

Il s'agit bien sûr de la méthode la plus utilisée, puisqu'il « suffit » de résoudre l'équation $u(x) = 0$.

Exemple

Soit $u : \begin{cases} \mathbb{R}[X] & \longrightarrow & \mathbb{R}[X] \\ P & \longmapsto & P(X+1) - P(X)\,. \end{cases}$

Montrer que u est linéaire et déterminer son noyau.

- *Le fait que u est un endomorphisme de $\mathbb{R}[X]$ a déjà été démontré dans un exemple page 243.*
- *Un polynôme P appartient à $\operatorname{Ker} u$ si et seulement si $P(X+1) = P(X)$. Cette relation implique, par récurrence immédiate :*
$$\forall\, n \in \mathbb{N},\ P(0) = P(1) = \cdots = P(n)\,.$$
Le polynôme $Q = P - P(0)$ s'annule donc sur $\mathbb{N}$; puisqu'il possède une infinité de racines, c'est le polynôme nul, et P est constant.
Réciproquement si P est un polynôme constant on a $u(P) = 0$ donc :
$$\operatorname{Ker}(u) = \mathbb{K}_0[X] \text{ (ensemble des polynômes constants).}$$
✓ *On notera que l'endomorphisme u est surjectif (puisque $\operatorname{Im} u = \mathbb{K}[X]$, voir exemple page 243) mais non injectif (puisque $\operatorname{Ker} u \neq \{0\}$) ; cet exemple montre bien que les propriétés énoncées en* **[S2.19]** *ne sont vraies qu'en dimension finie.*

SF2.41 Utiliser le théorème du rang

Dans certains cas, il est très facile de déterminer le rang d'une application linéaire. Le théorème du rang permet alors d'obtenir un renseignement sur le noyau.
L'exemple qui suit est vraiment très simple, mais il montre bien qu'une connaissance parfaite des différentes méthodes et théorèmes utilisés en algèbre linéaire peut permettre d'économiser bien des calculs.

Exemple

Déterminer le noyau de l'endomorphisme u de $\mathbb{R}^4$ dont la matrice dans la base canonique $\mathscr{B} = (e_1, e_2, e_3, e_4)$ est $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

En notant C_i pour $1 \leqslant i \leqslant 4$ les colonnes de la matrice, on remarque que :

– *les colonnes C_1, C_2 et C_3 sont linéairement indépendantes ;*

– *et $C_4 = C_1 - C_2 + C_3$.*

On en déduit que le rang de A, donc celui de u, est égal à 3. D'après le théorème du rang, le noyau de u est donc de dimension 1.
Or la dernière relation ci-dessus peut aussi s'écrire :

$$u(e_1) - u(e_2) + u(e_3) = u(e_4)\,,$$

soit $u(e_1 - e_2 + e_3 - e_4) = 0$. Ainsi le vecteur $e_1 - e_2 + e_3 - e_4$ est-il élément du noyau de u et, puisque la dimension du noyau est égale à 1 on a :

$$\operatorname{Ker} u = \operatorname{Vect}(e_1 - e_2 + e_3 - e_4)\,.$$

Exemple

Soit $a \in \mathbb{K}$, n un entier $\geqslant 3$ et u l'application qui à tout polynôme $P \in \mathbb{K}_n[X]$ associe le polynôme :

$$u(P) = (X - a)[P'(X) + P'(a)] - 2[P(X) - P(a)]\,.$$

Montrer que u est un endomorphisme de $\mathbb{K}_n[X]$. Déterminer son image et son noyau.

- *La linéarité de u découle de celle de la dérivation.*
- *Il faut ici penser à utiliser la base de $\mathbb{K}_n[X]$ formée des polynômes $(X-a)^k$ pour $k \in [\![0\,;n]\!]$ (c'est bien une base car formée de polynômes non nuls de degrés échelonnés de 0 à n).*
  Un calcul rapide (mais où il faut faire attention à distinguer les cas $k = 0$ et $k = 1$) donne :

  $$u(1) = u(X - a) = 0 \quad \text{et} \quad \forall\, k \in [\![2\,;n]\!],\ u\left((X - a)^k\right) = (k - 2)(X - a)^k\,.$$

  Puisque l'image de u est le sous-espace vectoriel engendré par les images des vecteurs de base, on obtient directement :

  $$\operatorname{Im} u = \operatorname{Vect}\left((X - a)^3, \ldots, (X - a)^n\right)$$

  (il s'agit des polynômes admettant a comme racine au moins triple).
  On en déduit que $\dim \operatorname{Im} u = (n + 1) - 3$ puis que $\dim \operatorname{Ker} u = 3$. Et puisque les vecteurs 1, $X - a$ et $(X - a)^2$ appartiennent au noyau, on a exactement :

  $$\operatorname{Ker} u = \operatorname{Vect}\left(1, X - a, (X - a)^2\right) = \mathbb{R}_2[X]\,.$$

  ✓ *En anticipant sur un prochain chapitre, on remarquera que la démonstration ci-dessus revenait à donner une base de vecteurs propres de u.*

  ✓ Cet exemple montre que la résolution d'un problème d'algèbre linéaire en dimension finie peut être grandement simplifiée lorsqu'on utilise une base adaptée.

À vous de jouer !

Exercice 2.1

SF2.4 **SF2.5** **SF2.26**

Soit E le $\mathbb{R}$-espace vectoriel des applications de classe $\mathscr{C}^1$ de $[0\,;1]$ dans $\mathbb{R}$, et :

$$F = \left\{ f \in E \;\middle|\; \int_0^1 f = 0 \text{ et } f(0) = f'(0) = 0 \right\}.$$

1) Montrer que F est un sous-espace vectoriel de E.

2) En trouver un supplémentaire.

Exercice 2.2

SF2.8 **SF2.12**

Soit $n \in \mathbb{N}^*$ et $(a_1, \ldots, a_n)$ des réels deux à deux distincts.

Démontrer que la famille des polynômes $(X - a_i)^n$ pour $i \in [\![1\,;n]\!]$ est libre.

Exercice 2.3

SF2.26 **SF2.35**

1) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et F, G deux sous-espaces vectoriels de E tels que $\dim(F) = \dim(G)$.
Montrer qu'il existe un sous-espace vectoriel H de E tel que $F \oplus H = G \oplus H = E$ (*on pourra procéder par récurrence sur* $n - \dim(F)$).

2) On ne suppose plus ici que E est de dimension finie.

a) Soient F et G deux sous-espaces vectoriels de E admettant un supplémentaire commun H, c'est-à-dire $E = F \oplus H = G \oplus H$.
Montrer que F et G sont isomorphes.

b) Soit $E = \mathbb{R}[X]$, $F = \text{Vect}\left(\left\{X^{2k} \;\middle|\; k \in \mathbb{N}\right\}\right)$ et $G = \text{Vect}\left(\left\{X^{2k} \;\middle|\; k \in \mathbb{N}^*\right\}\right)$.
Montrer que F et G sont isomorphes, mais qu'ils ne possèdent pas de supplémentaire commun.

Exercice 2.4

SF2.20 **SF2.36** **SF2.38** **SF2.39**

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, et u, $v \in \mathscr{L}(E)$.

1) Montrer : $\text{rg}(u + v) \leqslant \text{rg}\, u + \text{rg}\, v$.

2) Démontrer l'équivalence :

$$\text{rg}(u + v) = \text{rg}\, u + \text{rg}\, v \iff \left[\text{Im}\, u \cap \text{Im}\, v = \{0\} \quad \text{et} \quad E = \text{Ker}\, u + \text{Ker}\, v\right].$$

3) Montrer : $\text{rg}\, u + \text{rg}\, v - n \leqslant \text{rg}(u \circ v) \leqslant \min(\text{rg}\, u, \text{rg}\, v)$.

Exercice 2.5 **SF2.23**

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels, $f \in \mathscr{L}(E,F)$ et $g \in \mathscr{L}(F,G)$. Établir :

1) $\operatorname{Ker}(g \circ f) = \operatorname{Ker} f \iff \operatorname{Ker} g \cap \operatorname{Im} f = \{0\}$.

2) $\operatorname{Im}(g \circ f) = \operatorname{Im} g \iff \operatorname{Ker} g + \operatorname{Im} f = F$.

Exercice 2.6 **SF2.34**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $f \in \mathscr{L}(E)$.
Montrer qu'il existe un automorphisme $g \in \mathrm{GL}(E)$ tel que $f \circ g \circ f = f$.

Exercice 2.7 **SF2.23** **SF2.36** **SF2.40**

Soient p et q deux projecteurs tels que $p \circ q = q \circ p$.

1) Montrer que $p \circ q$ est un projecteur.

2) Montrer que $\operatorname{Ker}(p \circ q) = \operatorname{Ker} p + \operatorname{Ker} q$.

3) Montrer que $\operatorname{Im}(p \circ q) = \operatorname{Im} p \cap \operatorname{Im} q$.

Exercice 2.8 **SF2.27** **SF2.38**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$.
Soient u et v dans $\mathscr{L}(E)$ tels que $u + v = \mathrm{Id}_E$ et $\operatorname{rg} u + \operatorname{rg} v \leqslant n$.
Montrer que u et v sont des projecteurs.

Exercice 2.9 **SF2.34** **SF2.35**

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $u \in \mathscr{L}(E)$.

1) Montrer qu'il existe un projecteur $p \in \mathscr{L}(E)$ et un automorphisme $v \in \mathrm{GL}(E)$ tels que : $u = p \circ v$.

2) Montrer qu'il existe un projecteur $q \in \mathscr{L}(E)$ et un automorphisme $v \in \mathrm{GL}(E)$ tels que : $u = v \circ q$.

Thème 3 - Calcul matriciel

Comme le précédent, ce thème est consacré principalement à des révisions, mais il est indispensable de bien maîtriser ces notions avant d'aborder le chapitre sur la réduction.

■ Comment effectuer des calculs sur les matrices ?

SF3.1 Utiliser les formules

La première idée qui vient lorsqu'il s'agit de faire des calculs sur les matrices est d'utiliser les formules du cours donnant la somme et le produit de deux matrices. Cependant, ces formules mal utilisées peuvent conduire à des calculs inextricables ! Il s'agit donc d'une méthode à utiliser avec parcimonie et réflexion, ou bien lorsque la taille de la matrice est petite.

Exemple

Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$. Montrer que : $\operatorname{tr}({}^tAA) = 0 \implies A = 0_{n,p}$.

Si $A = (a_{k\ell})_{\substack{1\leqslant k\leqslant n\\ 1\leqslant \ell\leqslant p}}$, *le terme d'indice* (i,j) *de* ${}^tAA$ *pour* $(i,j) \in [\![1\,;p]\!]^2$ *est, par la formule du produit matriciel :*

$$\left({}^tAA\right)_{ij} = \sum_{k=1}^{n} \left({}^tA\right)_{ik} a_{kj} = \sum_{k=1}^{n} a_{ki}a_{kj}\,,$$

donc :

$$\operatorname{tr}\left({}^tAA\right) = \sum_{i=1}^{p} \left({}^tAA\right)_{ii} = \sum_{i=1}^{p}\sum_{k=1}^{n} a_{ki}^2\,.$$

Les termes a_{ki}^2 *étant des réels positifs, on a finalement :*

$$\operatorname{tr}\left({}^tAA\right) = 0 \implies \forall\,(k,i) \in [\![1\,;n]\!] \times [\![1\,;p]\!],\ a_{ki} = 0 \implies A = 0_{n,p}\,.$$

✓ Dans le cas $p = n$, cette propriété est utilisée lors de la définition du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{K})$ (voir **[S5.1]**).

Exemple

Montrer que si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ commute avec une matrice diagonale $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ où les λ_i sont *distincts*, alors A est diagonale.

Avec les notations habituelles pour les termes génériques des matrices, la propriété $AD = DA$ *s'écrit :*

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ (AD)_{ij} = (DA)_{ij}\,,$$

soit, par la formule du produit matriciel :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \sum_{k=1}^{n} a_{ik}d_{kj} = \sum_{k=1}^{n} d_{ik}a_{kj}\,.$$

Mais $d_{ik} = 0$ si $k \neq i$ et $d_{ii} = \lambda_i$ donc dans les sommes précédentes il ne reste qu'un terme et on obtient :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ a_{ij}\lambda_j = \lambda_i a_{ij} \quad \text{soit } (\lambda_i - \lambda_j)a_{ij} = 0\,.$$

Puisque $\lambda_i \neq \lambda_j$ pour $i \neq j$, on en déduit $a_{ij} = 0$ pour $i \neq j$, donc A est diagonale.

SF3.2 Utiliser les matrices de la base canonique

Lorsqu'une équation est linéaire, elle est vérifiée pour toute matrice si et seulement si elle l'est pour les matrices de la base canonique. Plutôt que de faire les calculs dans le cas général, il suffit donc dans ce cas de les faire avec ces matrices.

Exemple

Montrer que les matrices carrées qui commutent avec toutes les autres sont les matrices scalaires.

Dire qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ commute avec toutes les autres matrices carrées équivaut à dire qu'elle commute avec toutes les matrices E_{ij} de la base canonique. Or, pour $(i,j) \in [\![1\,;n]\!]^2$:

$$AE_{ij} = \left(\sum_{1\leqslant k,\ell\leqslant n} a_{k\ell}E_{kl}\right)E_{ij} = \sum_{1\leqslant k,\ell\leqslant n} a_{k\ell}E_{k\ell}E_{ij}$$

$$= \sum_{1\leqslant k,\ell\leqslant n} a_{k\ell}\delta_{\ell i}E_{kj} = \sum_{k=1}^{n} a_{ki}E_{kj}\,,$$

et de même : $E_{ij}A = \sum_{1\leqslant k,\ell\leqslant n} a_{k\ell}E_{ij}E_{k\ell} = \sum_{1\leqslant k,\ell\leqslant n} a_{k\ell}\delta_{jk}E_{i\ell} = \sum_{\ell=1}^{n} a_{j\ell}E_{i\ell}\,.$

La propriété de l'énoncé est donc équivalente à :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \sum_{k=1}^{n} a_{ki}E_{kj} = \sum_{\ell=1}^{n} a_{j\ell}E_{i\ell}\,.$$

Les matrices E_{ij} formant une famille libre, l'égalité ci-dessus n'est possible que si les coefficients des mêmes matrices dans les deux sommes sont égaux ; en considérant le coefficient de E_{ij} on obtient $a_{ii} = a_{jj}$; en considérant les coefficients de E_{kj} pour tout $k \neq i$ on obtient $a_{ki} = 0$.
Cela étant vrai pour tout $(i,j) \in [\![1\,;n]\!]^2$, la matrice A est scalaire.

✓ Nous vous conseillons de refaire soigneusement ces calculs afin de vous entraîner à manipuler des sommes contenant le symbole de Kronecker.

Exemple

Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$ telles que :

$$\forall\,M \in \mathcal{M}_n(\mathbb{K}),\ AMB = 0_n\,.$$

Montrer que A ou B est nulle.

Là encore, par linéarité, la propriété de l'énoncé est équivalente à :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ AE_{ij}B = 0_n\,.$$

Un calcul semblable à celui de l'exemple précédent, que nous vous conseillons de détailler, montre que :

$$\forall\,(i,j)\in[\![1\,;n]\!]^2,\ AE_{ij}B=\sum_{1\leqslant k,\ell\leqslant n}a_{ki}b_{j\ell}E_{k\ell}\,.$$

La famille $(E_{k\ell})$ étant libre, on obtient :

$$\forall\,(i,j,k,\ell)\in[\![1\,;n]\!]^4,\ a_{ki}b_{j\ell}=0\,.$$

Si la matrice A (par exemple) n'est pas nulle, il existe (au moins) un couple (k,i) tel que $a_{ki}\neq 0$; la relation ci-dessus donne alors $b_{j\ell}=0$ pour tous j, ℓ, c'est-à-dire $B=0_n$.

SF3.3 Faire des calculs par blocs

Dans certains cas, faire des calculs par blocs peut simplifier un peu les calculs. Ces cas sont aisément reconnaissables.

Exemple

Déterminer le commutant de la matrice $A=\begin{pmatrix}1&2&0&0\\2&1&0&0\\0&0&1&-2\\0&0&-1&1\end{pmatrix}\in\mathcal{M}_4(\mathbb{R})$.

Une petite astuce consiste ici à remarquer que M commute avec A si et seulement si M commute avec $A'=A-\mathrm{I}_4$; l'intérêt est que la matrice A' comporte plus de coefficients nuls, ce qui simplifie notablement les calculs. On cherche donc les matrices $M\in\mathcal{M}_4(\mathbb{R})$ telles que $A'M=MA'$. Il est naturel de chercher M sous la forme $M=\begin{bmatrix}M_1&M_2\\M_3&M_4\end{bmatrix}$ où les M_i sont des matrices de $\mathcal{M}_2(\mathbb{R})$. On est donc amené à résoudre :

$$\begin{bmatrix}A_1&0_2\\0_2&A_4\end{bmatrix}\times\begin{bmatrix}M_1&M_2\\M_3&M_4\end{bmatrix}=\begin{bmatrix}M_1&M_2\\M_3&M_4\end{bmatrix}\times\begin{bmatrix}A_1&0_2\\0_2&A_4\end{bmatrix}$$

avec $A_1=\begin{pmatrix}0&2\\2&0\end{pmatrix}$ et $A_4=\begin{pmatrix}0&-2\\-1&0\end{pmatrix}$, ce qui équivaut à :

$$A_1M_1=M_1A_1\ ,\ A_1M_2=M_2A_4\ ,\ A_4M_3=M_3A_1\ \ \text{et}\ \ A_4M_4=M_4A_4\,.$$

La résolution (qui est un peu pénible mais rapide) de quatre systèmes à quatre inconnues nous donne les matrices de la forme :

$$M=\begin{pmatrix}a&b&0&0\\b&a&0&0\\0&0&c&2d\\0&0&d&c\end{pmatrix}\quad\text{avec }(a,b,c,d)\in\mathbb{R}^4\,.$$

SF3.4 Utiliser la structure de $\mathcal{M}_n(\mathbb{K})$

Il ne faut pas oublier que $\mathcal{M}_n(\mathbb{K})$ possède une structure d'algèbre (*non commutative*) pour les 3 lois usuelles : on peut donc faire des calculs directement sur les matrices en utilisant les propriétés de ces lois.

Exemple

Soient A, $B \in \mathcal{M}_n(\mathbb{K})$ telles que :

$$A+B \in \mathrm{GL}_n(\mathbb{K}) \quad , \quad ABA = A^2B \quad \text{et} \quad BAB = B^2A\,.$$

Montrer que $AB = BA$.

On calcule :

$A(AB-BA) = A^2B - ABA = 0_n$ *et* $B(AB-BA) = BAB - B^2A = 0_n$.

En additionnant ces deux égalités on obtient : $(A+B)(AB-BA) = 0_n$.
$A+B$ *étant inversible, on en déduit* $AB - BA = 0_n$.

SF3.5 Utiliser les applications linéaires

Il s'agit certainement de la méthode la plus efficace dans beaucoup de cas. Il ne faut en effet pas oublier que les matrices servent avant tout à étudier des problèmes linéaires ; on a donc tout intérêt à considérer une matrice comme matrice d'une certaine application linéaire (entre des espaces vectoriels à préciser, munis de bases à préciser).

Exemple

Soient A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$ telles que : $\forall\, M \in \mathcal{M}_n(\mathbb{K})$, $AMB = 0_n$.
Montrer que A ou B est nulle.

Cet exemple a déjà été traité de façon purement calculatoire page 250 ; nous vous invitons à comparer les deux méthodes.
Notons respectivement a *et* b *les endomorphismes de* $\mathbb{R}^n$ *canoniquement associés aux matrices* A *et* B. *L'hypothèse* $AMB = 0_n$ *pour toute matrice* $M \in \mathcal{M}_n(\mathbb{K})$ *est donc équivalente à :*

$$\forall\, u \in \mathscr{L}(\mathbb{R}^n),\ a \circ u \circ b = 0_{\mathscr{L}(E)}\,.$$

Si b *n'est pas l'endomorphisme nul, son image contient au moins un vecteur non nul* $e' = b(e)$. *Pour tout vecteur* x *de* $\mathbb{R}^n$ *on peut trouver un endomorphisme* u *de* $\mathbb{R}^n$ *tel que* $u(e') = x$ *: il suffit pour cela de compléter le système libre* $\{e'\}$ *en une base de* $\mathbb{R}^n$ *et de définir* u *sur cette base, en utilisant la caractérisation d'une application linéaire (***[S2.18]***).*

On aura alors $a(x) = a \circ u(e') = a \circ u \circ b(e) = 0$, *et comme cela est vrai pour tout* $x \in \mathbb{R}^n$, *on a* $a = 0_{\mathscr{L}(E)}$.

■ Comment calculer les puissances d'une matrice ?

SF3.6 Procéder par récurrence

Le calcul des puissances A^p $(p \in \mathbb{N})$ d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ peut se faire par récurrence sur p principalement dans deux cas :

– si la formule donnant A^p en fonction de p vous est donnée ; il ne reste plus alors qu'à la vérifier ;

– si le calcul des premières puissances de A a fait apparaître une forme régulière ; on peut alors conjecturer la forme générale donnant A^p, puis la démontrer rigoureusement par récurrence.

Il est clair que si la matrice donnée n'a pas une forme « sympathique », cette dernière méthode n'a guère de chance d'aboutir…

Exemple

Soit J la matrice carrée d'ordre n dont tous les coefficients sont égaux à 1. Calculer J^p pour $p \in \mathbb{N}^*$.

La formule du produit matriciel montre que le coefficient d'indice (i,j) de la matrice J^2 est égal à :

$$\sum_{k=1}^{n} J_{ik}J_{kj} = \sum_{k=1}^{n} 1 = n\,,$$

donc $J^2 = nJ$.
À partir de cette relation, il est immédiat de démontrer par récurrence que $J^p = n^{p-1}J$ pour tout $p \in \mathbb{N}^*$.

SF3.7 Utiliser la formule du binôme

Cette méthode, détaillée dans les exemples ci-après, est rapide mais ne s'applique que dans peu de cas. On utilise le résultat suivant.

Si A et B sont des matrices de $\mathcal{M}_n(\mathbb{K})$ *telles que* $AB = BA$, alors :

$$\forall p \in \mathbb{N},\ (A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{p-k}\,.$$

✓ Attention ! cette formule tombe en défaut si A et B ne commutent pas.

Exemple

Soit $A = \begin{pmatrix} a & b & \dots & b \\ b & a & \dots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & \dots & b & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$.

Calculer A^p pour $p \in \mathbb{N}$, et éventuellement A^{-1} et A^p pour $p \in \mathbb{Z}$.

- *On introduit ici la matrice J carrée d'ordre n dont tous les éléments sont égaux à 1, déjà vue dans l'exemple précédent. On a ainsi $A = (a-b)\mathrm{I}_n + bJ$. On a vu que $J^k = n^{k-1}J$ pour tout $k \in \mathbb{N}^*$, donc, d'après la formule du binôme,*

$$\begin{aligned} A^p &= (a-b)^p \mathrm{I}_n + \sum_{k=1}^{p} \binom{p}{k} (a-b)^{p-k} b^k J^k \\ &= (a-b)^p \mathrm{I}_n + \left(\sum_{k=1}^{p} \binom{p}{k} (a-b)^{p-k} b^k n^{k-1} \right) J\,. \end{aligned}$$

Or :

$$\sum_{k=1}^{p}\binom{p}{k}(a-b)^{p-k}b^k n^{k-1} = \frac{1}{n}\sum_{k=1}^{p}\binom{p}{k}(a-b)^{p-k}b^k n^k$$
$$= \frac{1}{n}\left(\sum_{k=0}^{p}\binom{p}{k}(a-b)^{p-k}(bn)^k - (a-b)^p\right)$$
$$= \frac{1}{n}\big((a+(n-1)b)^p - (a-b)^p\big)\,,$$

donc $A^p = (a-b)^p \mathrm{I}_n + \frac{1}{n}\big((a+(n-1)b)^p - (a-b)^p\big)J \quad (*)$.

- *Le sous-espace vectoriel* $\mathcal{A}$ *de* $\mathcal{M}_n(\mathbb{R})$ *engendré par* I_n *et* J *(c'est-à-dire l'ensemble des matrices de la forme* $\lambda \mathrm{I}_n + \mu J$*) est aussi une sous-algèbre de* $\mathcal{M}_n(\mathbb{R})$ *(car* $J^2 \in \mathcal{A}$*). On utilise alors le lemme suivant.*

  Lemme : si $\mathcal{A}$ est une sous-algèbre de $\mathcal{M}_n(\mathbb{R})$ et si $A \in \mathcal{A}$ est inversible, alors $A^{-1} \in \mathcal{A}$.

  Démontrons-le.

  Puisque, si $M \in \mathcal{A}$, $AM \in \mathcal{A}$ *(*$\mathcal{A}$ *algèbre), on peut considérer l'application* $f : \begin{cases} \mathcal{A} & \longrightarrow & \mathcal{A} \\ M & \longmapsto & AM\,. \end{cases}$

  f *est un endomorphisme de l'espace vectoriel* $\mathcal{A}$*. Puisque* A *est inversible, on a* $\operatorname{Ker} f = \{0\}$*, donc* f *est injective. Puisque* $\mathcal{A}$ *est de dimension finie,* f *est bijective, donc* I_n *(qui appartient bien à la sous-algèbre* $\mathcal{A}$*) possède un antécédent par* f *dans* $\mathcal{A}$*. Cet antécédent, c'est forcément* A^{-1}*, et par suite,* $A^{-1} \in \mathcal{A}$.

Pour $A = (a-b)\mathrm{I}_n + bJ$*, on cherche donc son inverse éventuel sous la forme* $A' = a'\mathrm{I}_n + b'J$*. En utilisant le fait que la famille* (I,J) *est libre, la relation* $AA' = \mathrm{I}_n$ *conduit au système :*

$$\begin{cases} (a-b)a' = 1 \\ a'b + (a-b)b' + nbb' = 0 \end{cases}$$

qui admet une solution (a',b') *si et seulement si* $a-b \neq 0$ *et* $a+(n-1)b \neq 0$*, et dans ce cas* $a' = \dfrac{1}{a-b}$ *et* $b' = \dfrac{-b}{(a-b)\big(a+(n-1)b\big)}$*. On a donc trouvé* $A^{-1} = a'\mathrm{I}_n + b'J$.

- *On remarque alors que ces valeurs sont exactement celles obtenues pour* $p = -1$ *dans* $(*)$*. Cela suggère que la formule* $(*)$ *reste vraie pour* $p < 0$ *dans le cas où* A *est inversible (c'est-à-dire* $(a-b)\big(a+(n-1)b\big) \neq 0$*). Pour démontrer cela, on peut procéder par récurrence descendante sur* p *ou bien simplement vérifier l'égalité :*

$$\Big[(a-b)^p \mathrm{I}_n + \frac{1}{n}\big((a+(n-1)b)^p - (a-b)^p\big)J\Big]$$
$$\times \Big[(a-b)^{-p}\mathrm{I}_n + \frac{1}{n}\big((a+(n-1)b)^{-p} - (a-b)^{-p}\big)J\Big] = \mathrm{I}_n\,,$$

calcul que nous vous laissons le soin de faire...

Exemple

Soit $A = \begin{pmatrix} 1 & a & 0 & \dots & 0 \\ 0 & 1 & a & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \dots & \ddots & a \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$.

Calculer A^p pour $p \in \mathbb{N}$. Calculer A^{-1} et A^p pour $p \in \mathbb{Z}$.

- *On a* $A = \mathrm{I}_n + aN$ *avec* $N = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$.

  N est la matrice, dans la base canonique $(e_1, \dots, e_n)$, d'un endomorphisme u de $\mathbb{R}^n$ tel que $u(e_1) = 0$ et $u(e_i) = e_{i-1}$ pour $i \in [\![2\,;n]\!]$. On a alors $u^2(e_1) = u^2(e_2) = 0$ et $u^2(e_i) = e_{i-2}$ pour $i \in [\![3\,;n]\!]$, ce qui montre que :

  $$N^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & 0 & & \ddots & 0 \\ & & & & 0 \end{pmatrix}.$$

  On peut calculer ainsi les puissances successives de N jusqu'à :

  $$N^{n-1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ & \ddots & \ddots & \ddots & 0 \\ & & \ddots & \ddots & \vdots \\ & 0 & & \ddots & 0 \\ & & & & 0 \end{pmatrix}, \text{ puis } N^n = 0_n\,.$$

  La formule du binôme s'applique puisque I_n et N commutent :

  $$A^p = (\mathrm{I}_n + aN)^p = \sum_{k=0}^{p} \binom{p}{k} a^k N^k\,.$$

  Puisque $N^k = 0$ pour $k \geqslant n$, on obtient :

  $$A^p = \begin{pmatrix} 1 & \binom{p}{1}a & \binom{p}{2}a^2 & \dots & \dots & \binom{p}{n-1}a^{n-1} \\ 0 & 1 & \binom{p}{1}a & \ddots & & \binom{p}{n-2}a^{n-2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \binom{p}{2}a^2 \\ \vdots & & & \ddots & \ddots & \binom{p}{1}a \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix},$$

  cette formule étant vraie pour tout entier $p \in \mathbb{N}$ avec la convention habituelle $\binom{p}{k} = 0$ si $k > p$.

- *On peut alors étendre la définition des coefficients binomiaux en posant :*
  - *si* $p<0$ *:* $\binom{p}{k}=\dfrac{p(p-1)\dots(p-k+1)}{k!}$ *si* $k\in\mathbb{N}^*$ *et* $\binom{p}{0}=1$ *;*
  - *si* $k<0$, $\binom{p}{k}=0$ *pour tout* $p\in\mathbb{Z}$.

  On vérifie alors facilement que la formule du triangle de Pascal :
  $$\binom{p}{k}=\binom{p-1}{k}+\binom{p-1}{k-1}$$
  est encore valable pour tout $(p,k)\in\mathbb{Z}^2$*. Il en résulte que, en notant* $M_p=\sum_{k=0}^{+\infty}\binom{p}{k}(aN)^k$ *pour tout* $p\in\mathbb{Z}$*, on aura encore la relation :* $M_p\times(\mathrm{I}_n+aN)=M_{p+1}$*, la démonstration étant la même que celle de la formule du binôme.*

  Cela permet de montrer par récurrence descendante sur p *que la relation* $A^p=(\mathrm{I}_n+aN)^p=M_p$ *reste vraie pour tout entier* $p<0$.

- *En particulier, pour* $p=-1$*, on a* $\binom{-1}{k}=\dfrac{(-1)(-2)\cdots(-k)}{k!}=(-1)^k$ *pour tout* $k\in\mathbb{N}^*$*, ce qui donne la matrice : :*
$$A^{-1}=\begin{pmatrix}1&-a&a^2&\dots&\dots&(-1)^{n-1}a^{n-1}\\0&1&-a&\ddots&&(-1)^{n-2}a^{n-2}\\\vdots&\ddots&\ddots&\ddots&\ddots&\vdots\\\vdots&&\ddots&\ddots&\ddots&a^2\\\vdots&&&\ddots&\ddots&-a\\0&\dots&\dots&\dots&0&1\end{pmatrix}.$$

  ✓ *Le calcul direct de* A^{-1} *pouvait se faire très rapidement en appliquant la méthode exposée en* **[SF3.21]**.

SF3.8 Utiliser les applications linéaires

Comme nous l'avons déjà mentionné, oublier les simples calculs et considérer une matrice comme celle d'une application linéaire bien choisie conduit à des méthodes en général plus performantes. C'est ce que nous avons fait pour calculer les puissances de N dans l'exemple précédent. Voici un autre exemple.

Exemple

Calculer les puissances de la matrice $A=\begin{pmatrix}0&1&1&\dots&1\\1&0&0&\dots&0\\1&&&&\\\vdots&&0&&\\1&0&\dots&\dots&0\end{pmatrix}$.

En notant $(e_1,\dots,e_n)$ *la base canonique de* $\mathbb{R}^n$ *et* u *l'endomorphisme canoniquement associé à* A*, on a par définition :*

$$u(e_1) = e_2 + \cdots + e_n \quad \text{et} \quad u(e_i) = e_1 \text{ pour tout } i \in [\![2\,;n]\!]\,.$$

On en tire :

$$u^2(e_1) = \sum_{i=2}^{n} u(e_i) = (n-1)e_1 \quad \text{et} \quad \forall\, i \in [\![2\,;n]\!],\ u^2(e_i) = e_2 + \cdots + e_n$$

puis $u^3(e_1) = (n-1)(e_2 + \cdots + e_n)$ *et* $\forall\, i \in [\![2\,;n]\!],\ u^3(e_i) = (n-1)e_1$ *.*
On en déduit $u^3 = (n-1)u$ *(puisque les deux applications linéaires coïncident sur une base), d'où* $A^3 = (n-1)A$*, et ensuite, par récurrence :*

$$\forall\, p \in \mathbb{N}^*,\ A^{2p} = (n-1)^p A^2 \ \text{ et } \ \forall\, p \in \mathbb{N},\ A^{2p+1} = (n-1)^p A\,.$$

SF3.9 Utiliser la réduction

Les méthodes les plus générales pour calculer les puissances d'une matrice sont celles utilisant la réduction, en particulier l'utilisation d'un polynôme annulateur **[SF4.16]** et la diagonalisation **[SF4.17]** qui sont détaillées à partir de la page 309.

✓ Sauf quelques cas particuliers comme dans les exemples précédents, ce sont ces méthodes, exposées dans le chapitre suivant, qu'il faut privilégier.

■ Comment calculer le rang d'une matrice ?

SF3.10 Utiliser la définition

Le rang d'une matrice est la dimension du sous-espace vectoriel engendré par ses vecteurs colonnes. On peut utiliser cette propriété lorsque l'on remarque que les colonnes d'une matrice sont des combinaisons linéaires de certains vecteurs fixés.

Exemple

Calculer le rang de la matrice $A = \begin{pmatrix} 1 & 2^2 & \dots & n^2 \\ 2^2 & 3^2 & \dots & (n+1)^2 \\ \vdots & \vdots & \ddots & \vdots \\ n^2 & (n+1)^2 & \dots & (2n-1)^2 \end{pmatrix}$.

Le terme d'indice (i,j) *de* A *est égal à :*

$$a_{ij} = (i+j-1)^2 = (i-1)^2 + 2j(i-1) + j^2\,,$$

donc on pourra écrire chaque colonne C_j *sous la forme :*

$$C_j = U + 2j.V + j^2.W\,,$$

$$\text{avec : } U = \begin{pmatrix} 0 \\ 1^2 \\ 2^2 \\ \vdots \\ (n-1)^2 \end{pmatrix},\ V = \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ n-1 \end{pmatrix} \ \text{ et } \ W = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

- *Si $n = 1$, $A = (1)$ est de rang 1 et si $n = 2$, $A = \begin{pmatrix} 1 & 4 \\ 4 & 9 \end{pmatrix}$ est de rang 2.*
- *Si $n \geqslant 3$, toutes les colonnes de A sont dans le sous-espace vectoriel de $\mathcal{M}_{n,1}(\mathbb{K})$ engendré par U, V et W, donc le rang de A est inférieur ou égal à 3. Puisque la matrice extraite formée des 3 premières lignes et colonnes de A est $A' = \begin{pmatrix} 1 & 4 & 9 \\ 4 & 9 & 16 \\ 9 & 16 & 25 \end{pmatrix}$ qui est inversible ($\det A' = -8$), le rang de A est supérieur ou égal à 3.*
  En conclusion, le rang de A est égal à 3 si $n \geqslant 3$.

SF3.11 Utiliser les opérations élémentaires

Le rang d'une matrice étant invariant par les opérations élémentaires sur les lignes ou sur les colonnes exposées en **[S3.13]**, on peut calculer le rang d'une matrice en la rendant équivalente (par lignes ou/et par colonnes) à une matrice plus simple (triangulaire en général).

✓ Pour utiliser cette méthode, nul besoin de mettre la matrice sous forme échelonnée réduite complète. Bien souvent, quelques opérations élémentaires bien choisies suffisent pour conclure.

Exemple

Déterminer le rang de $A = \begin{pmatrix} a & 0 & \dots & \dots & 0 & b \\ b & a & 0 & \dots & \dots & 0 \\ 0 & b & a & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & b & a & 0 \\ 0 & 0 & \dots & 0 & b & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$.

On distingue deux cas.

- *Si $a = 0$: dans ce cas, si $b = 0$ la matrice A est nulle (donc de rang 0), et si $b \neq 0$, A est inversible (donc de rang n) puisque, en permutant les colonnes on obtient $b\mathrm{I}_n$.*
- *Si $a \neq 0$: les opérations $L_i \longleftarrow L_i - \frac{b}{a} L_{i-1}$ pour i variant de 2 à n donnent la matrice équivalente :*

$$A' = \begin{pmatrix} a & 0 & \dots & \dots & 0 & b \\ 0 & a & 0 & \dots & \dots & -\frac{b^2}{a} \\ 0 & 0 & a & 0 & \dots & \frac{b^3}{a^2} \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & a & (-1)^{n-2}\frac{b^{n-1}}{a^{n-2}} \\ 0 & 0 & \dots & 0 & 0 & a + (-1)^{n-1}\frac{b^n}{a^{n-1}} \end{pmatrix}$$

Cette matrice A' est triangulaire supérieure ; puisque l'on a déjà supposé $a \neq 0$, le rang de cette matrice est égal à n sauf si $a + (-1)^{n-1}\dfrac{b^n}{a^{n-1}} = 0$, soit $a^n = (-b)^n$ ou encore $\dfrac{-b}{a}$ racine n^e de l'unité, et dans ce dernier cas le rang est égal à $n-1$.

SF3.12 Utiliser une application linéaire

Si la matrice A est la matrice d'une certaine application linéaire u, on sait que le rang de A est égal au rang de u. On peut alors essayer de déterminer plutôt le rang de u, par exemple en utilisant le théorème du rang.

Exemple

En l'interprétant comme la matrice d'un endomorphisme de $\mathbb{R}_n[X]$, déterminer

le rang de la matrice $A = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & 0 \\ n & 1 & 2 & 0 & \dots & 0 \\ 0 & n-1 & 1 & 3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 2 & 1 & n \\ 0 & 0 & \dots & 0 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{n+1}(\mathbb{R})$.

Soit u l'endomorphisme de $\mathbb{R}_n[X]$ dont A est la matrice dans la base canonique. On a donc :

$$\begin{cases} u(1) = 1 + nX,\ u(X^n) = nX^{n-1} + X^n, \\ u(X^k) = kX^{k-1} + X^k + (n-k)X^{k+1} \text{ si } k \in [\![1\,;n-1]\!]\,. \end{cases}$$

On en déduit que pour tout $P \in \mathbb{R}_n[X]$ on a :

$$u(P) = (1-X^2)P' + (nX+1)P\,,$$

car cette égalité est vérifiée pour les vecteurs de la base canonique.

Cherchons le noyau de u :

$$u(P) = 0 \iff (X^2-1)P' = (nX+1)P.$$

La résolution de l'équation différentielle $y' = \dfrac{nx+1}{x^2-1}y$ sur l'intervalle $]1\,;+\infty[$ donne :

$$y = C\exp\left(\int \frac{nx+1}{x^2-1}\,\mathrm{d}x\right) = C\exp\left(\int\left(\frac{\frac{n+1}{2}}{x-1} + \frac{\frac{n-1}{2}}{x+1}\right)\mathrm{d}x\right)$$
$$= C(x-1)^{\frac{n+1}{2}}(x+1)^{\frac{n-1}{2}} \quad (C \text{ constante réelle})\,.$$

Les solutions trouvées sont polynomiales si et seulement si n est impair. En conclusion, le noyau de u est une droite si n est impair et est réduit à $\{0\}$ si n est pair.

On en déduit, puisque $\dim \operatorname{Ker} u + \operatorname{rg} u = \dim \mathbb{R}_n[X] = n+1$, *que le rang de A est égal à $n+1$ si n est pair et à n si n est impair.*

SF3.13 Utiliser des matrices extraites

On sait que le rang d'une matrice est aussi la taille de la plus grande matrice carrée inversible que l'on peut extraire de A ; il en découle que, si l'on peut extraire d'une matrice A une matrice carrée de rang r, alors $\operatorname{rg} A \geqslant r$. Cela permet souvent d'avoir une inégalité intéressante.
Nous avons déjà utilisé ce résultat lors de l'exemple traité en **[SF3.10]**. En voici un autre, qui montre l'intérêt de cette méthode.

Exemple

Déterminer le rang de la matrice $A = \begin{pmatrix} 2 & a & a & a \\ a & 2 & a & a \\ a & a & 2 & a \\ a & a & a & 2 \end{pmatrix}$ selon les valeurs de a.

Dans ce genre d'exercices, on commence déjà par traiter le cas où la matrice est inversible. Pour cela, le plus simple ici est de calculer son déterminant. Ce calcul est classique et sera exposé dans un exemple en **[SF3.36]**. *On obtient :* $\det A = (2+3a)(2-a)^3$.

- *Si* $a \neq -\dfrac{2}{3}$ *et* $a \neq 2$, $\det A \neq 0$ *: A est inversible, donc* $\operatorname{rg} A = 4$.
- *Si* $a = 2$, *la matrice est évidemment de rang* 1.
- *Si* $a = -\dfrac{2}{3}$, *on a* $\operatorname{rg} A \leqslant 3$. *Mais l'on remarque que la matrice extraite* $\begin{pmatrix} 2 & -2/3 & -2/3 \\ -2/3 & 2 & -2/3 \\ -2/3 & -2/3 & 2 \end{pmatrix}$ *est inversible (son déterminant vaut* $\dfrac{128}{27}$*). Donc le rang de A est supérieur ou égal à* 3 *et, finalement,* $\operatorname{rg} A = 3$.

SF3.14 Utiliser l'équivalence avec J_r

Nous avons vu en **[S3.11]** qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est de rang r si et seulement si il existe deux matrices P, $Q \in \mathrm{GL}_n(\mathbb{K})$ telles que :

$$A = PJ_rQ \quad \text{où } J_r = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix} \text{ (si } r = 0,\ J_r = 0_n\text{)}.$$

L'exemple ci-dessous montre comment utiliser ce résultat.

Exemple

Soient $A \in \mathcal{M}_p(\mathbb{K})$ et $B \in \mathcal{M}_q(\mathbb{K})$ deux matrices carrées d'ordres p, $q \in \mathbb{N}^*$. On forme la matrice par blocs $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathcal{M}_{p+q}(\mathbb{K})$.

Montrer que $\operatorname{rg} M = \operatorname{rg} A + \operatorname{rg} B$.

Notons $r = \operatorname{rg} A$ *et* $r' = \operatorname{rg} B$. *En utilisant le résultat rappelé ci-dessus, on obtient qu'il existe* $(P,Q) \in \mathrm{GL}_p(\mathbb{K})^2$ *et* $(P',Q') \in \mathrm{GL}_q(\mathbb{K})^2$ *telles que* $A = PJ_rQ$ *et* $B = P'J_{r'}Q'$, *où* J_r *est d'ordre p et* $J_{r'}$ *d'ordre q.*
On aura alors, en effectuant un produit par blocs :

$$M = \begin{bmatrix} P & 0 \\ 0 & P' \end{bmatrix} \times \begin{bmatrix} J_r & 0 \\ 0 & J_{r'} \end{bmatrix} \times \begin{bmatrix} Q & 0 \\ 0 & Q' \end{bmatrix}.$$

Les matrices $\begin{bmatrix} P & 0 \\ 0 & P' \end{bmatrix}$ *et* $\begin{bmatrix} Q & 0 \\ 0 & Q' \end{bmatrix}$ *étant inversibles, on en déduit que le rang de* M *est égal au rang de la matrice* $\begin{bmatrix} J_r & 0 \\ 0 & J_{r'} \end{bmatrix}$*, qui est une matrice diagonale avec* $r + r'$ *termes égaux à* 1 *sur la diagonale les autres étant nuls, donc* $\operatorname{rg} M = r + r' = \operatorname{rg} A + \operatorname{rg} B$*.*

✓ *Pour justifier précisément que la matrice* $\begin{bmatrix} P & 0 \\ 0 & P' \end{bmatrix}$ *est inversible, on peut faire son produit avec* $\begin{bmatrix} P^{-1} & 0 \\ 0 & P'^{-1} \end{bmatrix}$*, qui donne la matrice identité* I_{p+q}*, ou noter que son déterminant est égal à* $\det P \times \det P'$*, donc est non nul.*

SF3.15 Utiliser des calculs par blocs

Pour trouver le rang d'une matrice écrite par blocs, il est naturel de penser à faire des opérations élémentaires sur les blocs (nous en verrons des exemples en **[SF3.41]**), ou d'utiliser des produits par blocs.

Exemple

Soient n et p deux entiers naturels non nuls, $A \in \mathrm{GL}_n(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$, $C \in \mathcal{M}_{p,n}(\mathbb{K})$ et $D \in \mathcal{M}_p(\mathbb{K})$.

Soit enfin $M \in \mathcal{M}_{n+p}(\mathbb{K})$ donnée par : $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

Montrer que : $\operatorname{rg} M = n \iff D = CA^{-1}B$.

On effectue le produit par blocs :

$$\begin{bmatrix} \mathrm{I}_n & 0 \\ CA^{-1} & -\mathrm{I}_p \end{bmatrix} \times \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & CA^{-1}B - D \end{bmatrix}.$$

Or la matrice $\begin{bmatrix} \mathrm{I}_n & 0 \\ CA^{-1} & -\mathrm{I}_p \end{bmatrix}$ *est inversible (car, par exemple, elle est triangulaire inférieure à éléments diagonaux non nuls), donc puisque l'on ne change pas le rang d'une matrice en la multipliant par une matrice inversible,*

$$\operatorname{rg} M = \operatorname{rg} \begin{bmatrix} A & B \\ 0 & CA^{-1}B - D \end{bmatrix}.$$

Or, d'après l'exemple de **[SF3.14]** *page 260, le rang de cette matrice est égal à* $\operatorname{rg} A + \operatorname{rg}(CA^{-1}B - D)$*, d'où le résultat.*

■ Comment calculer l'inverse d'une matrice ?

SF3.16 Utiliser la comatrice

La formule $A^{-1} = \dfrac{1}{\det A}\,{}^{\mathrm{t}}\mathrm{com}(A)$ n'est utilisable directement que pour des matrices de petite taille.

Par exemple, pour une matrice de taille 2, elle permet de montrer que, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible, alors : $A^{-1} = \dfrac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Elle reste également utilisable pour calculer assez rapidement l'inverse d'une matrice 3×3.

Exemple

Calculer l'inverse de la matrice $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 3 \end{pmatrix}$.

$$\det P = \begin{vmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & -1 & 3 \end{vmatrix} = 6 \text{ et } \operatorname{com}(P) = \begin{pmatrix} \begin{vmatrix} 0 & 2 \\ -1 & 3 \end{vmatrix} & -\begin{vmatrix} -1 & 2 \\ 0 & 3 \end{vmatrix} & \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ -1 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 0 & -1 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ -1 & 0 \end{vmatrix} \end{pmatrix},$$

ce qui conduit à :

$$\operatorname{com}(P) = \begin{pmatrix} 2 & 3 & 1 \\ -4 & 3 & 1 \\ 2 & -3 & 1 \end{pmatrix} \quad \text{puis} \quad P^{-1} = \frac{1}{6}\begin{pmatrix} 2 & -4 & 2 \\ 3 & 3 & -3 \\ 1 & 1 & 1 \end{pmatrix}.$$

En dehors de ces cas, la comatrice est surtout utilisée dans certains problèmes théoriques.

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{Z})$. Déterminer une condition nécessaire et suffisante pour que A soit inversible dans l'anneau $\mathcal{M}_n(\mathbb{Z})$.

- *Si $A \in \mathcal{M}_n(\mathbb{Z})$ est inversible dans l'anneau $\mathcal{M}_n(\mathbb{Z})$, il existe $B \in \mathcal{M}_n(\mathbb{Z})$ telle que $AB = \mathrm{I}_n$. On a alors* $\det A \times \det B = 1$.
  Mais $\det A$ *et* $\det B$ *sont des entiers (comme produit d'entiers, par la formule qui donne le déterminant). On a donc* $\det A = \pm 1$.
- *Réciproquement, si $A \in \mathcal{M}_n(\mathbb{Z})$ est telle que* $\det A = \pm 1$, *elle est inversible dans $\mathcal{M}_n(\mathbb{R})$, et son inverse est donnée par la formule rappelée ci-dessus, qui s'écrit ici : $A^{-1} = \pm {}^{\mathrm{t}}\operatorname{com}(A)$. Mais la comatrice de A est elle aussi à coefficients entiers, puisque ses termes sont des déterminants extraits de A. La matrice A^{-1} est donc à coefficients entiers, et A est inversible dans l'anneau $\mathcal{M}_n(\mathbb{Z})$.*
- *En conclusion :*

  $$A \in \mathcal{M}_n(\mathbb{Z}) \text{ est inversible dans l'anneau } \mathcal{M}_n(\mathbb{Z}) \iff \det A = \pm 1\,.$$

SF3.17 Résoudre le système $Y = AX$

Cette méthode est la traduction dans le cas matriciel de la méthode de base pour trouver la réciproque d'une application ; en effet, si f est une application bijective d'un ensemble A sur un ensemble B on a, par définition :

$$\forall\,(x,y) \in A \times B,\ y = f(x) \iff x = f^{-1}(y)\,,$$

autrement dit, si f permet d'exprimer y en fonction de x, il suffit d'exprimer x en fonction de y pour obtenir f^{-1}.

Si A est une matrice carrée inversible et X, Y des vecteurs colonnes, on a ainsi :

$$Y = AX \iff X = A^{-1}Y\,,$$

donc il suffit de résoudre le système $Y = AX$ pour obtenir A^{-1}.

✓ Le fait que ce système a une solution X unique permet de prouver en même temps que A est inversible.

Exemple

Déterminer l'inverse de $A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \dots & n \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$ $(a_{ij} = \min(i,j))$.

Soient $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ *et* $Y = {}^{\mathrm{t}}\begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}$ *deux éléments de* $\mathcal{M}_{n,1}(\mathbb{R})$. *Pour résoudre le système* $Y = AX$, *écrivons sa matrice complète :*

$$\left(\begin{array}{ccccccc|c} 1 & 1 & 1 & \dots & \dots & \dots & 1 & y_1 \\ 1 & 2 & 2 & \dots & \dots & \dots & 2 & y_2 \\ 1 & 2 & 3 & \dots & \dots & \dots & 3 & y_3 \\ & & & & \vdots & & & \vdots \\ 1 & 2 & 3 & \dots & n-2 & n-2 & n-2 & y_{n-2} \\ 1 & 2 & 3 & \dots & \dots & n-1 & n-1 & y_{n-1} \\ 1 & 2 & 3 & \dots & \dots & n-1 & n & y_n \end{array}\right).$$

Par les opérations élémentaires $L_i \leftarrow L_i - L_{i-1}$ *pour* i *variant de* n *à* 2, *le système est équivalent à celui de matrice :*

$$\left(\begin{array}{ccccccc|c} 1 & 1 & 1 & \dots & \dots & \dots & 1 & y_1 \\ 0 & 1 & 1 & \dots & \dots & \dots & 1 & y_2 - y_1 \\ 0 & 0 & 1 & \dots & \dots & \dots & 1 & y_3 - y_2 \\ & & & & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 & 1 & y_{n-2} - y_{n-3} \\ 0 & 0 & 0 & \dots & 0 & 1 & 1 & y_{n-1} - y_{n-2} \\ 0 & 0 & 0 & \dots & \dots & 0 & 1 & y_n - y_{n-1} \end{array}\right),$$

puis par les opérations élémentaires $L_i \leftarrow L_i - L_{i-1}$ *pour* i *variant de* 1 *à* $n-1$, *on obtient :*

$$\begin{cases} x_1 &= 2y_1 - y_2 \\ x_2 &= -y_1 + 2y_2 - y_3 \\ \vdots & \vdots \\ x_{n-1} &= -y_{n-2} + 2y_{n-1} - y_n \\ x_n &= y_n - y_{n-1} \end{cases}.$$

Le système ayant une solution unique, la matrice A est inversible, et les calculs précédents donnent :

$$A^{-1} = \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & 0 & & \\ 0 & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & 0 \\ & & \ddots & -1 & 2 & -1 \\ & & & 0 & -1 & 1 \end{pmatrix}.$$

SF3.18 Interpréter la matrice comme une matrice de passage

Si A est la matrice de passage d'une base $\mathscr{B} = (e_1, \ldots, e_n)$ à une base $\mathscr{B}' = (e'_1, \ldots, e'_n)$, on sait que A^{-1} est alors la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$. Ainsi, si A permet d'exprimer les e'_j en fonction des e_i, il suffit d'exprimer les e_i en fonction des e'_j pour obtenir A^{-1}.
On est donc encore ramené à inverser un système, dont les inconnues sont ici des vecteurs. Cette méthode est donc tout à fait similaire à la précédente, la différence étant que l'on raisonne ici sur les colonnes plutôt que sur les lignes.

Exemple

Déterminer l'inverse de $A = \begin{pmatrix} 1 & a & \ldots & a \\ & 1 & \ddots & \vdots \\ & & \ddots & a \\ 0 & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$.

On remarque déjà que A est bien inversible, car elle est triangulaire supérieure à éléments diagonaux non nuls.
Si l'on note $\mathscr{B} = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$, A est la matrice de passage de cette base à une certaine base $\mathscr{B}' = (e'_1, \ldots, e'_n)$ telle que :

$$e'_1 = e_1,\ e'_2 = e_2 + ae_1,\ e'_3 = e_3 + ae_2 + ae_1,\ \ldots,\ e'_j = e_j + ae_{j-1} + \cdots + ae_1, \ldots$$

On résout ce système :

$$e_1 = e'_1\ ,\ e_2 = e'_2 - ae'_1\ ,\ e_3 = e'_3 - ae'_2 - a(1-a)e'_1\ , \ldots,$$
$$e_j = e'_j - ae'_{j-1} - a(1-a)e'_{j-2} - \cdots - a(1-a)^{j-1}e'_1,\ \ldots$$

donc :

$$A^{-1} = \begin{pmatrix} 1 & -a & -a(1-a) & \ldots & \ldots & -a(1-a)^{n-2} \\ & 1 & -a & \ddots & & -a(1-a)^{n-3} \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & -a & -a(1-a) \\ & 0 & & & 1 & -a \\ & & & & 0 & 1 \end{pmatrix}.$$

SF3.19 Utiliser la méthode du pivot complète

Nous avons vu en **[S3.13]** que si l'on applique à la matrice identité la même suite d'opérations élémentaires que celle qui a permis d'obtenir la forme échelonnée réduite d'une matrice A, alors on obtenait la matrice A^{-1}.

Disons-le nettement : *cette méthode ne présente strictement aucun intérêt* ! En effet, les opérations élémentaires qui permettent de passer de la matrice A à la matrice identité peuvent aussi très bien se faire sur les lignes du système $Y = AX$, ce qui permettra d'inverser le système et de « lire » directement A^{-1} (c'est ce que nous avons fait pour l'exemple page 263). Alors pourquoi s'embêter à faire aussi les mêmes opérations sur la matrice identité ? Vous multipliez ainsi les calculs et donc les risques d'erreur…
De plus, obtenir un système échelonné suffit en général pour le résoudre ; nul besoin lorsqu'on utilise cette méthode de réduire en plus la matrice pour mettre des pivots égaux à 1 ni des 0 au dessus. Il y a donc en général moins de calculs à faire.

On oublie donc cette pseudo méthode pourtant si prisée de certains d'entre vous !

SF3.20 Utiliser une application linéaire associée

Cette méthode consiste à utiliser la propriété suivante : si A est la matrice dans une certaine base d'un automorphisme u, alors A^{-1} est la matrice de u^{-1} dans cette même base.

✓ Si l'on veut seulement démontrer que A est inversible, c'est-à-dire que u est un automorphisme, il suffit de démontrer que $\operatorname{Ker} u = \{0\}$, ce qui équivaut à démontrer que la seule solution du système $AX = 0$ est $X = 0$.

Exemple

Déterminer l'inverse de $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ où $a_{ij} = \binom{j-1}{i-1}$ (avec la convention habituelle $\binom{j-1}{i-1} = 0$ si $i > j$).

Par la formule du binôme on a, pour tout $j \in [\![1\,;n]\!]$ *:*

$$(X+1)^{j-1} = \sum_{i=0}^{j-1} \binom{j-1}{i} X^i = \sum_{i=1}^{n} \binom{j-1}{i-1} X^{i-1}.$$

La matrice A est donc la matrice de passage de la base canonique $(1,X,\dots,X^{n-1})$ de $\mathbb{R}_{n-1}[X]$ à la base formée des $(X+1)^{j-1}$ pour $1 \leqslant j \leqslant n$ ou aussi la matrice dans la base canonique de $\mathbb{R}_{n-1}[X]$ de l'endomorphisme $u\colon P(X) \mapsto P(X+1)$.

u est un automorphisme de $\mathbb{R}_{n-1}[X]$ et u^{-1} est l'endomorphisme qui à tout polynôme $P(X)$ associe $P(X-1)$. La matrice de u^{-1} dans la base canonique est celle dont les colonnes sont les coordonnées de $(X-1)^{j-1}$ dans cette base ; la matrice A^{-1} a donc pour coefficients les $(-1)^{j-i}\binom{j-1}{i-1}$.

SF3.21 Faire l'analogie avec les séries entières

Nous allons déjà exposer cette méthode dans un cas particulier.
Supposons que $A \in \mathcal{M}_n(\mathbb{K})$ s'écrive sous la forme $A = \mathrm{I}_n + N$, où I_n est la matrice identité et N une matrice nilpotente, c'est-à-dire $N^n = 0_n$. On cherche A^{-1}, ce qui fait penser à l'inverse (dans $\mathbb{R}$) de $1 + x$. Or, on connaît le développement en série entière :

$$\forall\, x \in]-1\,;1[,\ \frac{1}{1+x} = \sum_{k=0}^{+\infty} (-1)^k x^k\,,$$

ce qui peut aussi s'écrire :

$$\forall\, x \in]-1\,;1[,\ (1+x)\left(\sum_{k=0}^{+\infty} (-1)^k x^k\right) = 1\,.$$

On a donc aussi envie d'écrire (pour l'instant, sans justification) :

$$(\mathrm{I}_n + N)\left(\sum_{k=0}^{+\infty} (-1)^k N^k\right) = \mathrm{I}_n\,,$$

et puisque $N^k = 0_n$ pour $k \geqslant n$, cette relation s'écrit aussi :

$$(\mathrm{I}_n + N)(\mathrm{I}_n - N + N^2 - \cdots + (-1)^{n-1} N^{n-1}) = \mathrm{I}_n\,.$$

Il est alors facile de vérifier proprement cette dernière égalité en développant simplement le produit.
Cela prouve que $\mathrm{I}_n + N$ est inversible (on ne le savait pas a priori) et que :

$$(\mathrm{I}_n + N)^{-1} = \mathrm{I}_n - N + N^2 - \cdots + (-1)^{n-1} N^{n-1}\,.$$

Exemple

Soit n un entier, $n \geqslant 2$. Inverser la matrice $A \in \mathcal{M}_n(\mathbb{R})$ dont les coefficients a_{ij} valent 0 si $i > j$, 1 si $i = j$ et -1 sinon.

On commence par écrire la matrice sous forme développée :

$$A = \begin{pmatrix} 1 & -1 & \dots & -1 \\ & 1 & \ddots & \vdots \\ & & \ddots & -1 \\ 0 & & & 1 \end{pmatrix}.$$

On voit déjà que A est triangulaire supérieure, avec des coefficients non nuls sur la diagonale. Elle est donc inversible.
La méthode consistant à résoudre le système $Y = AX$, à condition de faire des opérations convenables sur les lignes, s'applique très bien ici (et donc aussi la méthode du pivot) ; vous pouvez vous entraîner à résoudre l'exercice de cette façon.
Pour utiliser la méthode précédente, nous allons introduire la matrice :

$$J = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & 0 \\ & & \ddots & \ddots \\ & 0 & & \ddots & 1 \\ & & & & 0 \end{pmatrix}.$$

On a alors :

$$J^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & 0 & & \ddots & 0 \\ & & & & 0 \end{pmatrix}, \; J^{n-1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ & \ddots & \ddots & \ddots & 0 \\ & & \ddots & \ddots & \vdots \\ & 0 & & \ddots & 0 \\ & & & & 0 \end{pmatrix}, \; J^n = 0_n$$

(ces résultats ont été établis dans un exemple de **[SF3.7]** *page 255).*

On peut donc écrire : $A = \mathrm{I}_n - J - \cdots - J^{n-1} = 2\mathrm{I}_n - \sum_{k=0}^{+\infty} J^k$.

On pense alors naturellement à utiliser les séries entières. Si x *est un réel de* $]-1\,;1[$*, on a :*

$$S = 2 - \sum_{k=0}^{+\infty} x^k = 2 - \frac{1}{1-x} = \frac{1-2x}{1-x}\,,$$

donc pour $|x| < \frac{1}{2}$ *:*

$$\frac{1}{S} = \frac{1-x}{1-2x} = \frac{1-2x+x}{1-2x} = 1 + \frac{x}{1-2x} = 1 + x\sum_{k=0}^{+\infty}(2x)^k = 1 + \sum_{k=1}^{+\infty} 2^{k-1}x^k\,.$$

On a donc, pour $|x| < \frac{1}{2}$ *:*

$$\left(2 - \sum_{k=0}^{+\infty} x^k\right)\left(1 + \sum_{k=1}^{+\infty} 2^{k-1}x^k\right) = 1\,.$$

On reconnaît ici un produit de Cauchy de deux séries absolument convergentes ; la formule donnant ce produit étant la même dans toute algèbre normée, l'égalité reste vraie en l'appliquant à une matrice M *sous réserve que les séries soient absolument convergentes.*
C'est le cas ici puisque la matrice J *est nilpotente ; ainsi :*

$$\left(2\mathrm{I}_n - \sum_{k=0}^{+\infty} J^k\right)\left(\mathrm{I}_n + \sum_{k=1}^{+\infty} 2^{k-1}J^k\right) = \mathrm{I}_n\,.$$

Donc l'inverse de A *est la matrice* $\mathrm{I}_n + J + 2J^2 + \cdots + 2^{n-2}J^{n-1}$*, soit :*

$$A^{-1} = \begin{pmatrix} 1 & 2 & 4 & \dots & 2^{n-2} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 4 \\ & 0 & & \ddots & 2 \\ & & & & 1 \end{pmatrix}.$$

SF3.22 Utiliser un polynôme annulateur

Si l'on connaît un polynôme annulateur P de la matrice A, de la forme $P = \sum_{k=0}^{d} a_k X^k$, avec $a_0 \neq 0$ (condition indispensable), alors on a :

$$a_0 I = -a_1 A - a_2 A^2 - \ldots - a_d A^d = A(-a_1 I - a_2 A - \ldots - a_d A^{d-1})\,,$$

ce qui prouve que A est inversible et a pour inverse

$$A^{-1} = \frac{1}{a_0}(-a_1 I - a_2 A - \ldots - a_d A^{d-1}).$$

On peut donc calculer A^{-1} à l'aide des premières puissances de A.
La méthode générale pour trouver un polynôme annulateur d'une matrice sera exposée dans le chapitre concernant la réduction.

Exemple

Soit $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Trouver une relation entre A^2, A et I_2 et en déduire A^{-1}.

On vérifie facilement que $A^2 - 4A + 3\mathrm{I}_2 = 0_2$.
On en déduit que $\mathrm{I}_2 = \frac{1}{3}A(4\mathrm{I}_2 - A)$ *donc* A *est inversible et :*

$$A^{-1} = \frac{1}{3}(4\mathrm{I}_2 - A) = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

✓ Plus généralement, pour toute matrice carrée d'ordre 2, on a la relation :

$$A^2 - (\operatorname{tr} A)A + (\det A)\mathrm{I}_2 = 0_2\,.$$

SF3.23 Utiliser la réduction

Si la matrice A est diagonalisable et inversible, on a $A = PDP^{-1}$ avec $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ (les λ_i étant non nuls) et P inversible, d'où $A^{-1} = PD^{-1}P^{-1}$, avec $D^{-1} = \operatorname{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}\right)$.

Cette méthode est en fait peu pratique puisque :

– il faut déjà que A soit diagonalisable ;

– il faut déterminer les valeurs propres puis les vecteurs propres ;

– et enfin, il faut calculer P^{-1}.

Cette méthode n'est donc pas à utiliser d'emblée, uniquement si des questions précédentes ont fait diagonaliser A.

SF3.24 Faire des calculs par blocs

Rien de très original ici : il s'agit juste de rappeler que pour chercher l'inverse d'une matrice écrite par blocs, il est naturel d'utiliser des produits par blocs !

Exemple

Soient A, B, C, $D \in \mathcal{M}_n(\mathbb{K})$ et $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{K})$.

On suppose M, A et D inversibles. Calculer M^{-1}.

Tout naturellement, on cherche M^{-1} *(qui existe par hypothèse) sous la forme* $M^{-1} = \begin{bmatrix} X & Y \\ Z & T \end{bmatrix}$, *avec* X, Y, Z, $T \in \mathcal{M}_n(\mathbb{K})$.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \times \begin{bmatrix} X & Y \\ Z & T \end{bmatrix} = \mathrm{I}_{2n} \iff \begin{cases} AX + BZ = \mathrm{I}_n \\ AY + BT = 0_n \\ CX + DZ = 0_n \\ CY + DT = \mathrm{I}_n \end{cases}$$
$$\iff \begin{cases} AX + BZ = \mathrm{I}_n \\ Y = -A^{-1}BT \\ Z = -D^{-1}CX \\ CY + DT = \mathrm{I}_n \end{cases}$$

puisque A et D sont inversibles par hypothèse. La première et la dernière équations s'écrivent alors :

$$(A - BD^{-1}C)X = \mathrm{I}_n \quad \text{et} \quad (-CA^{-1}B + D)T = \mathrm{I}_n\,.$$

Puisque l'on sait que ce système a une solution (car M inversible), il en résulte que les matrices $A - BD^{-1}C$ et $-CA^{-1}B + D$ sont forcément inversibles, et l'on aura finalement :

$$M^{-1} = \begin{bmatrix} \left(A - BD^{-1}C\right)^{-1} & -A^{-1}B\left(D - CA^{-1}B\right)^{-1} \\ -D^{-1}C\left(A - BD^{-1}C\right)^{-1} & \left(D - CA^{-1}B\right)^{-1} \end{bmatrix}.$$

✓ Remarquez dans cet exercice (et dans d'autres) comme nous avons utilisé indifféremment l'inverse à gauche et/ou l'inverse à droite d'une matrice. Cela est suffisant, en vertu d'un théorème rappelé en **[S3.4]**.

SF3.25 Reconnaître une matrice orthogonale

Nous citons juste cette méthode très particulière afin qu'elle ne soit pas oubliée : si M est une matrice orthogonale, $M^{-1} = {}^{\mathrm{t}}M$. Pour plus de détails, nous vous renvoyons à **[S6.1]**.

Exemple

Déterminer l'inverse de $A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{pmatrix}$.

Cette matrice n'est pas orthogonale, mais l'on remarque que ses colonnes sont orthogonales deux à deux et ont même norme (pour le produit scalaire canonique dans $\mathbb{R}^3$).
La matrice $\frac{1}{3}A$ est donc, elle, une matrice orthogonale : ses vecteurs colonnes forment une base orthonormale de $\mathbb{R}^3$. Donc $\left(\frac{1}{3}A\right)^{-1} = {}^{\mathrm{t}}\left(\frac{1}{3}A\right)$, d'où l'on tire : $A^{-1} = \frac{1}{9}\,{}^{\mathrm{t}}A$.

■ Comment déterminer la matrice d'une application linéaire ?

Notons déjà que le titre de cette problématique est ambigu : il ne faut jamais parler de LA matrice d'une application linéaire, sans préciser dans quelle(s) base(s) elle est écrite. Ce choix est d'ailleurs fondamental, car travailler dans une « bonne » base peut simplifier grandement l'expression des matrices considérées.

SF3.26 Revenir à la définition

Il s'agit de la méthode la plus courante : la matrice d'une application linéaire dans des bases $\mathscr{B}$ et $\mathscr{B}'$ est obtenue en calculant les coordonnées dans la base $\mathscr{B}'$ des vecteurs de $u(\mathscr{B})$.

Exemple

Soit $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ une matrice donnée dans $\mathcal{M}_n(\mathbb{K})$. On considère l'endomorphisme u de $\mathcal{M}_n(\mathbb{K})$ défini par :

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}),\ u(M) = MA\,.$$

Déterminer la matrice de u dans la base canonique (E_{ij}) de $\mathcal{M}_n(\mathbb{K})$, les E_{ij} étant ordonnés selon l'ordre lexicographique.

On commence donc par chercher l'image par u des matrices E_{ij}. Nous avons déjà fait le calcul page 250 :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ u(E_{ij}) = E_{ij}A = \sum_{\ell=1}^{n} a_{j\ell}E_{i\ell}\,.$$

Lorsque l'on écrit les n^2 matrices E_{ij} dans l'ordre lexicographique, on obtient n blocs de n matrices, le i-ème bloc étant constitué des n matrices :

$$E_{i1}, E_{i2}, \ldots, E_{in}\,.$$

Pour tout j, $u(E_{ij})$ ne dépend que de ces n matrices et ses coordonnées sur ce bloc sont les $a_{j\ell}$ c'est-à-dire les éléments de la j^{e} ligne de A. La matrice de la restriction de u à $\operatorname{Vect}(E_{i1},E_{i2},\ldots,E_{in})$ sera donc égale à ${}^{\text{t}}A$ et au final, la matrice cherchée sera la matrice diagonale par blocs :

$$\begin{bmatrix} {}^{\text{t}}A & 0_n & \ldots & 0_n \\ 0_n & {}^{\text{t}}A & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_n \\ 0_n & \ldots & 0_n & {}^{\text{t}}A \end{bmatrix}.$$

SF3.27 Déterminer l'expression analytique

L'expression analytique d'une application linéaire u dans des bases $\mathscr{B}$ et $\mathscr{B}'$ est l'expression des coordonnées dans $\mathscr{B}'$ du vecteur $y = u(x)$ en fonction de celles du vecteur x dans $\mathscr{B}$; si l'on connaît cette expression, qui s'écrit matriciellement $Y = AX$ où A est la matrice de u dans les bases $\mathscr{B}$ et $\mathscr{B}'$, on en déduira A.

Exemple

Donner la matrice dans la base canonique de $\mathbb{R}^3$ de la projection p sur le plan P d'équation $x + y + z = 0$ parallèlement à la droite D d'équation $x = \dfrac{1}{2}y = \dfrac{1}{3}z$.

On remarque déjà que la droite D n'est pas incluse dans P ; donc P et D sont supplémentaires, et cela a bien un sens de considérer la projection sur P parallèlement à D.

Nous allons donc trouver l'expression analytique de p, c'est-à-dire les coordonnées (x',y',z') de l'image $p(u)$ par p d'un vecteur u de coordonnées (x,y,z).
Pour cela, on écrit simplement les deux conditions qui caractérisent la projection sur P parallèlement à D :

$$\forall\, u \in \mathbb{R}^3\ ,\ p(u) \in P \quad \text{et} \quad p(u) - u \in D\,.$$

La condition $p(u) \in P$ se traduit par $x' + y' + z' = 0$. La condition $p(u) - u \in D$ peut aussi s'écrire $p(u) - u$ colinéaire à un vecteur de base de D, par exemple $(1,2,3)$, soit encore :

$$\exists\, \lambda \in \mathbb{R} \text{ tel que } p(u) - u = \lambda.(1,2,3)\,.$$

Les coordonnées de $p(u) - u$ étant $(x' - x, y' - y, z' - z)$, on obtient :

$$x' - x = \lambda\ ;\ y' - y = 2\lambda\ ;\ z' - z = 3\lambda\,.$$

En additionnant ces trois égalités, compte tenu de $x'+y'+z' = 0$, on trouve :

$$\lambda = -\frac{1}{6}(x + y + z)\,,$$

d'où :
$$\begin{cases} x' = x + \lambda & = \dfrac{5}{6}x - \dfrac{1}{6}y - \dfrac{1}{6}z \\ y' = y + 2\lambda & = -\dfrac{1}{3}x + \dfrac{2}{3}y - \dfrac{1}{3}z \\ z' = z + 3\lambda & = -\dfrac{1}{2}x - \dfrac{1}{2}y + \dfrac{1}{2}z \end{cases}.$$

Les coefficients dans l'expression analytique de p donnent directement sa matrice cherchée (il suffit de lire !) :

$$A = \begin{pmatrix} \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

✓ La méthode présentée ci-dessus est certainement la plus rapide pour trouver la matrice d'une projection dans le cas général. Il faut donc la retenir.

SF3.28 Utiliser les formules de changement de base

Si l'on connaît la matrice d'une application linéaire dans certaines bases, les formules rappelées en **[S3.10]** peuvent permettre d'obtenir sa matrice dans des bases différentes.

Exemple

Donner la matrice dans la base canonique de $\mathbb{R}^3$ de la projection p sur le plan P d'équation $x + y + z = 0$ parallèlement à la droite D d'équation $x = \frac{1}{2}y = \frac{1}{3}z$.

Nous traitons ici le même exemple que précédemment, afin que vous puissiez comparer les deux méthodes.
Notons $\mathscr{B} = (e_1,e_2,e_3)$ la base canonique de $\mathbb{R}^3$, et A la matrice dans cette base de la projection cherchée.

Une base du plan P d'équation $x+y+z=0$ est, par exemple, formée des vecteurs $e'_1=(1,-1,0)$ et $e'_2=(1,0,-1)$. Une base de la droite D d'équation $x=\frac{1}{2}y=\frac{1}{3}z$ est par exemple $e'_3=(1,2,3)$.

Dans la base $\mathscr{B}'=(e'_1,e'_2,e'_3)$ formée de la réunion d'une base de P et d'une base de D, la matrice de la projection p est : $A'=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$ (puisque P est l'ensemble des vecteurs invariants par p et D est son noyau).

La matrice de passage Q de la base canonique $\mathscr{B}$ à la base $\mathscr{B}'$ est :

$$Q=\begin{pmatrix}1&1&1\\-1&0&2\\0&-1&3\end{pmatrix}.$$

D'après le cours, on a la relation $A'=Q^{-1}AQ$, donc la matrice cherchée est $A=QA'Q^{-1}$. Il reste à calculer Q^{-1}, ce que nous vous laissons le soin de faire à titre d'exercice, en utilisant l'une des méthodes vues auparavant.

On trouve : $Q^{-1}=\frac{1}{6}\begin{pmatrix}2&-4&2\\3&3&-3\\1&1&1\end{pmatrix}$.

Il ne reste plus qu'à faire le dernier calcul : $A=QA'Q^{-1}$ soit :

$$A=\frac{1}{6}\begin{pmatrix}1&1&1\\-1&0&2\\0&-1&3\end{pmatrix}\times\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\times\begin{pmatrix}2&-4&2\\3&3&-3\\1&1&1\end{pmatrix},$$

ce qui donne : $A=M_{\mathscr{B}}(p)=\frac{1}{6}\begin{pmatrix}5&-1&-1\\-2&4&-2\\-3&-3&3\end{pmatrix}$.

✓ Il est clair que cette méthode est bien moins performante que la précédente, puisqu'il a fallu calculer l'inverse et deux produits de matrices. Son avantage est cependant de pouvoir s'appliquer à d'autres endomorphismes que les projections, lorsque l'on connaît leurs propriétés dans une certaine base.

■ Comment montrer que deux matrices sont (ou pas) semblables ?

SF3.29 Utiliser des invariants de similitude

On sait que si deux matrices sont semblables, leurs trace, rang, déterminant, polynôme caractéristique, polynôme minimal… sont égaux. Mais la réciproque de cette propriété est fausse. On peut donc seulement l'utiliser pour montrer éventuellement que deux matrices ne sont pas semblables (ou pour vérifier certains calculs).

Exemple

Les matrices $A=\begin{pmatrix}1&1&1\\0&2&0\\0&0&2\end{pmatrix}$ et $B=\begin{pmatrix}1&0&1\\0&2&1\\0&0&2\end{pmatrix}$ sont-elles semblables ?

On remarque déjà que ces deux matrices ont même polynôme caractéristique $(X-1)(X-2)^2$*, mais cela ne permet pas de conclure.*
Cela fait cependant songer à utiliser les résultats sur la diagonalisation.
Plus précisément, 2 *est valeur propre double de* A *et puisque* $\mathrm{rg}(A-2\mathrm{I}_3)=1$*,* $\dim E_2(A)=\dim\mathrm{Ker}(A-2\mathrm{I}_3)=2$*, et* A *est diagonalisable.*
Mais $\mathrm{rg}(B-2\mathrm{I}_3)=2$ *donc* $\dim E_2(B)=1$ *et* B *n'est donc pas diagonalisable.*
A *et* B *ne sont donc pas semblables (sinon, par transitivité,* B *serait semblable à une matrice diagonale donc diagonalisable).*

SF3.30 Utiliser le calcul matriciel

Disons-le nettement : vouloir montrer à l'aide des formules sur le produit matriciel, que deux matrices sont semblables en écrivant $A'=P^{-1}AP$, même si l'on pense à transformer cette égalité en $PA'=AP$ pour éviter le calcul de P^{-1}, est peine perdue dès que la taille de la matrice est supérieure ou égale à 3 (c'est-à-dire presque tout le temps !). Cette méthode conduit en effet à résoudre un système de n^2 équations à n^2 inconnues...
Si l'utilisation de la relation $A'=P^{-1}AP$ n'est pas pratique pour le calcul, elle est cependant très utile voire incontournable dans des exercices plus théoriques.

Exemple

Montrer que si deux matrices de $\mathcal{M}_n(\mathbb{R})$ sont semblables dans $\mathcal{M}_n(\mathbb{C})$, elles sont aussi semblables dans $\mathcal{M}_n(\mathbb{R})$.

Soient $A,\ B\in\mathcal{M}_n(\mathbb{R})$*, semblables dans* $\mathcal{M}_n(\mathbb{C})$*, c'est-à-dire qu'il existe* $P\in\mathrm{GL}_n(\mathbb{C})$ *telle que* $B=P^{-1}AP$*, ce que l'on réécrit :* $PB=AP$*.*
En séparant les parties réelles et imaginaires des coefficients de P*, on peut écrire :*

$$P=P_1+\mathrm{i}P_2 \quad \textit{avec} \quad P_1,\ P_2\in\mathcal{M}_n(\mathbb{R})\,,$$

d'où $(P_1+\mathrm{i}P_2)B=A(P_1+\mathrm{i}P_2)$*, et les matrices* A *et* B *étant à coefficients réels, on aura, en identifiant les parties réelles et imaginaires des coefficients :*

$$P_1B=AP_1 \quad (1) \qquad \textit{et} \qquad P_2B=AP_2 \quad (2)\,.$$

La fonction $f:t\mapsto\det(P_1+tP_2)$ *est une fonction polynôme de degré* $\leqslant n$ *à coefficients réels. Puisque* P *est inversible, on a* $f(\mathrm{i})\neq 0$*, donc* f *n'est pas la fonction nulle. Il existe donc un réel* λ *tel que* $f(\lambda)\neq 0$ *(car si* f *était nulle sur* $\mathbb{R}$ *elle serait identiquement nulle comme fonction polynôme ayant une infinité de racines), c'est-à-dire* $P_1+\lambda P_2$ *inversible dans* $\mathcal{M}_n(\mathbb{R})$*.*
En combinant alors les relations (1) *et* (2)*, on obtient :*

$$(P_1+\lambda P_2)B=A(P_1+\lambda P_2) \quad \textit{soit} \quad B=(P_1+\lambda P_2)^{-1}A(P_1+\lambda P_2)\,,$$

ce qui démontre la proposition.

SF3.31 Utiliser les applications linéaires

L'interprétation d'une matrice carrée comme matrice d'un endomorphisme est bien sûr LA méthode à utiliser. Rappelons que dire que deux matrices sont semblables signifie qu'elles représentent le même endomorphisme dans des bases différentes ; l'écriture matricielle $A'=P^{-1}AP$ n'en est qu'une conséquence.

Pour montrer que A et A' de $\mathcal{M}_n(\mathbb{K})$ sont semblables, on pourra donc considérer l'endomorphisme u de $\mathbb{K}^n$ canoniquement associé à A ; il ne reste plus qu'à trouver une base $\mathscr{B}'$ de $\mathbb{K}^n$ telle que $A' = M_{\mathscr{B}'}(u)$.

Exemple

Pour tout $a \in \mathbb{R}$, on considère les matrices :

$$A_a = \begin{pmatrix} 4-a & 1 & -1 \\ -6 & -1-a & 2 \\ 2 & 1 & 1-a \end{pmatrix} \text{ et } B_a = \begin{pmatrix} 1-a & 1 & 0 \\ 0 & 1-a & 0 \\ 0 & 0 & 2-a \end{pmatrix}.$$

Montrer que A_a et B_a sont semblables pour tout $a \in \mathbb{R}$.

- *On remarque déjà que l'on peut écrire $A_a = A_0 - a\mathrm{I}_3$ et $B_a = B_0 - a\mathrm{I}_3$; puisque $P^{-1}(B_0 - a\mathrm{I}_3)P = P^{-1}B_0P - a\mathrm{I}_3$, A_a et B_a sont semblables pour tout a si et seulement si A_0 et B_0 le sont.*
  On est donc ramené à un problème sans paramètre.
- *Soit u l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique (e_1,e_2,e_3) est $A_0 = \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$.*

  On veut montrer qu'il existe une base (e'_1,e'_2,e'_3) de $\mathbb{R}^3$ telle que la matrice de u dans cette base soit égale à $B_0 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, ce qui équivaut à :

  $$u(e'_1) = e'_1 \ , \ u(e'_2) = e'_1 + e'_2 \ \text{ et } \ u(e'_3) = 2e'_3 \, .$$

  Notons que cela revient à trigonaliser la matrice A_0, et que la première et la dernière des relations précédentes ne sont rien d'autre que la recherche des vecteurs propres de A_0.
  - *Notons $e'_1 = (x,y,z)$ (coordonnées dans la base canonique). Alors :*

    $$u(e'_1) = e'_1 \iff \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
    $$\iff \begin{cases} 3x + y - z = 0 \\ -6x - 2y + 2z = 0 \\ 2x + y = 0 \end{cases} \iff \begin{cases} x = z \\ y = -2x \end{cases}.$$

    Le vecteur $e'_1 = (1,-2,1)$ convient.
  - *Notons $e'_2 = (x,y,z)$. Alors :*

    $$u(e'_2) = e'_1 + e'_2 \iff \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
    $$\iff \begin{cases} 3x + y - z = 1 \\ -6x - 2y + 2z = -2 \\ 2x + y = 1 \end{cases} \iff \begin{cases} x = z \\ y = -2x + 1 \end{cases}.$$

    Dans ce dernier calcul, on n'a pas oublié de vérifier que le système obtenu est bien compatible (sinon, A_0 et B_0 ne seraient pas semblables). On peut donc choisir le vecteur $e'_2 = (1,-1,1)$, qui vérifie ce système et est bien linéairement indépendant avec e'_1.

– *Notons enfin* $e'_3 = (x,y,z)$. *Alors :*

$$u(e'_3) = 2e'_3 \iff \begin{pmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\iff \begin{cases} 2x + y - z = 0 \\ -6x - 3y + 2z = 0 \\ 2x + y - z = 0 \end{cases} \iff \begin{cases} y = -2x \\ z = 0 \end{cases}.$$

Le vecteur $e'_3 = (1, -2, 0)$ *convient.*

- *On vérifie alors facilement que la famille* (e'_1, e'_2, e'_3) *est une base de* $\mathbb{R}^3$ *; la matrice de passage de la base canonique à cette nouvelle base est* $P = \begin{pmatrix} 1 & 1 & 1 \\ -2 & -1 & -2 \\ 1 & 1 & 0 \end{pmatrix}$, *et l'on a :* $B_a = P^{-1} A_a P$.

✓ Nous avons volontairement très détaillé l'exemple ci-dessus car nous n'en traiterons pas d'autre : c'est toujours la même chose !

■ Comment utiliser la trace ?

La définition de la trace d'une matrice **[S3.12]** est élémentaire. Mais ses propriétés sont plus importantes qu'on ne le croit, comme nous allons le voir dans les exemples qui suivent, et comme nous le verrons aussi dans le chapitre sur la réduction.

SF3.32 Utiliser la trace des matrices

Outre sa définition, la principale propriété de la trace qui est utilisée lors de calculs matriciels est sa linéarité.

Exemple

Soient A, $B \in \mathcal{M}_n(\mathbb{K})$, vérifiant : $AB - BA = A$.

Montrer que A n'est pas inversible.

On raisonne par l'absurde. Si A *était inversible, on aurait, en multipliant la relation de l'énoncé par* A^{-1} *à gauche :*

$$B - A^{-1}BA = \mathrm{I}_n .$$

On aurait alors :

$$\mathrm{tr}(B - A^{-1}BA) = \mathrm{tr}\, B - \mathrm{tr}(A^{-1}BA) = \mathrm{tr}\, \mathrm{I}_n = n ,$$

ce qui est impossible puisque $\mathrm{tr}(A^{-1}BA) = \mathrm{tr}(AA^{-1}B) = \mathrm{tr}\, B$.

✓ *On peut en fait montrer qu'une matrice* A *vérifiant la propriété de l'énoncé est forcément nilpotente (cf. exercice* **4.11***).*

Exemple

Soient A, $B \in \mathcal{M}_n(\mathbb{C})$, $A \neq 0_n$.

Résoudre l'équation : $X + \mathrm{tr}(X)A = B$, d'inconnue $X \in \mathcal{M}_n(\mathbb{C})$.

Supposons qu'il existe une solution $X \in \mathcal{M}_n(\mathbb{C})$. Alors, en prenant la trace des deux membres de l'égalité, on obtient (puisque tr *est une forme linéaire) :*

$$(1 + \operatorname{tr} A) \operatorname{tr} X = \operatorname{tr} B\,.$$

- *Si* $1 + \operatorname{tr} A \neq 0$, *on a nécessairement :* $\operatorname{tr} X = \dfrac{\operatorname{tr} B}{1 + \operatorname{tr} A}$, *d'où* $X = B - \dfrac{\operatorname{tr} B}{1 + \operatorname{tr} A} A$.

  Réciproquement, si $X = B - \dfrac{\operatorname{tr} B}{1 + \operatorname{tr} A} A$, *on vérifie aisément que* X *est bien solution de l'équation proposée. C'en est donc l'unique solution.*

- *Si* $\operatorname{tr} A = -1$, *deux autres cas sont à distinguer.*
  - *Si* $\operatorname{tr} B \neq 0$, *l'équation proposée n'a pas de solution.*
  - *Si* $\operatorname{tr} B = 0$, *on vérifie alors que* B *est solution. En posant* $Y = X - B$, *l'équation se ramène à* $Y = -\operatorname{tr}(Y) A$. Y *est donc de la forme* λA *avec* $\lambda \in \mathbb{K}$.

    Réciproquement, on vérifie aisément que toute matrice de ce type convient.
    En conclusion, dans ce cas, l'ensemble des solutions est l'ensemble des matrices de la forme $X = \lambda A + B$ *avec* $\lambda \in \mathbb{K}$.

SF3.33 Utiliser la trace des endomorphismes

Le principal intérêt de la trace d'une matrice est qu'il s'agit d'un invariant de similitude. C'est cette propriété qui permet de définir la trace d'un endomorphisme. Dans l'exemple ci-dessous, nous voyons une utilisation intéressante de la propriété : $\operatorname{tr} p = \operatorname{rg} p$ lorsque p est un projecteur.

Exemple

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.
Soit $(p_i)_{1 \leqslant i \leqslant k}$ une famille de projecteurs de E . Démontrer l'équivalence des propriétés suivantes :

(i) $\forall (i,j) \in [\![1\,;k]\!]^2,\ i \neq j \implies p_i \circ p_j = 0\,,$

(ii) $p = \displaystyle\sum_{i=1}^{k} p_i$ est un projecteur.

- (i) $\implies$ (ii)

  Si (i) *est vérifiée, alors :* $p^2 = \displaystyle\sum_{1 \leqslant i,j \leqslant k} p_i \circ p_j = \sum_{i=1}^{k} p_i^2 = \sum_{i=1}^{k} p_i = p\,,$

  donc p *est un projecteur.*

- (ii) $\implies$ (i)

  On a l'inclusion : $\operatorname{Im} p = \operatorname{Im}\left(\displaystyle\sum_{i=1}^{k} p_i\right) \subset \sum_{i=1}^{k} \operatorname{Im} p_i$, *donc*

  $$\dim \operatorname{Im} p \leqslant \dim\left(\sum_{i=1}^{k} \operatorname{Im} p_i\right) \leqslant \sum_{i=1}^{k} \dim(\operatorname{Im} p_i)\,,$$

ou encore : $\operatorname{rg} p \leqslant \sum_{i=1}^{k} \operatorname{rg} p_i$.

Si l'on suppose la propriété (ii) *vérifiée, puisque* p *et les* p_i *sont des projecteurs, on a* $\operatorname{rg} p = \operatorname{tr} p$ *et* $\operatorname{rg} p_i = \operatorname{tr} p_i$. *Or, par linéarité,* $\operatorname{tr} p = \sum_{i=1}^{k} \operatorname{tr} p_i$, *soit* $\operatorname{rg} p = \sum_{i=1}^{k} \operatorname{rg} p_i$. *Les inégalités précédentes sont donc des égalités. Il en résulte* $\dim\left(\sum_{i=1}^{k} \operatorname{Im} p_i\right) = \sum_{i=1}^{k} \dim(\operatorname{Im} p_i)$ *et* $\operatorname{Im} p = \sum_{i=1}^{k} \operatorname{Im} p_i$ *donc* $\operatorname{Im} p = \bigoplus_{i=1}^{k} \operatorname{Im} p_i$.

Pour tout $x \in E$ *et tout* $j \in [\![1\,;k]\!]$, $p_j(x) \in \operatorname{Im} p_j \subset \operatorname{Im} p$ *donc* $p[p_j(x)] = p_j(x)$ *(les vecteurs de* $\operatorname{Im} p$ *sont invariants par* p*).*

Ainsi $\sum_{i=1}^{k} p_i[p_j(x)] = p_j(x)$ *avec* $p_i[p_j(x)] \in \operatorname{Im} p_i$ *et* $p_j(x) \operatorname{Im} p_j$. *La somme des* $\operatorname{Im} p_i$ *étant directe, on en déduit* $p_i[p_j(x)] = 0$ *si* $i \neq j$, *soit la propriété* (i).

■ Comment calculer un déterminant ?

Ce n'est pas parce que ce paragraphe ne traite que de calcul et non de problèmes plus théoriques qu'il doit être négligé, au contraire. Justifions nos propos.

- La plupart d'entre vous rencontrent des difficultés pour mener un calcul intelligemment, rigoureusement, et avec méthode. Vous n'êtes pas une machine, donc il ne faut pas vous contenter d'aligner des pages de calcul ; il faut aussi réfléchir à chaque étape du calcul pour choisir la méthode qui rendra le résultat plus agréable, en utilisant des simplifications, des factorisations, en remarquant des symétries, des répétitions ou des analogies, en utilisant certaines transformations, en faisant attention aux valeurs particulières qui peuvent intervenir, en repérant des identités remarquables etc...
  Tout cela demande surtout de l'entraînement et de la persévérance.
- Les méthodes de calcul de déterminants sont relativement nombreuses et il faut les connaître toutes. Mais, il est rare qu'une seule méthode suffise, il faut souvent en appliquer plusieurs de suite. De plus, même deux déterminants qui se ressemblent peuvent conduire à des calculs très différents, aussi bien par la méthode que par la longueur. Un calcul de déterminant n'est jamais si facile !

Dans ce paragraphe, les corrigés de certains exemples ne seront pas aussi détaillés qu'à l'habitude ; nous nous contenterons d'indiquer les principales étapes des calculs, en vous laissant le soin de compléter vous-même les calculs intermédiaires. Essayez également de trouver d'autres méthodes que celles que nous proposons.
Vous remarquerez aussi que certains exemples apparaissent plusieurs fois : il est fréquent qu'il y ait plusieurs méthodes possibles pour calculer un déterminant !

SF3.34 Faire des combinaisons linéaires de lignes/colonnes

Il s'agit de la méthode de base : par des combinaisons linéaires judicieuses sur les rangées, qui ne changent pas le déterminant, on cherche à se ramener à une matrice triangulaire ; le déterminant sera alors le produit des éléments diagonaux.

Exemple

Calculer le déterminant d'ordre $n+1$: $\Delta(a_1,\dots,a_n) = \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & a_1 & 0 & \dots & 0 \\ 1 & 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & a_n \end{vmatrix}$.

Si les a_i sont non nuls, on fait l'opération $C_0 \leftarrow C_0 - \sum_{i=1}^{n} \frac{1}{a_i} C_i$, et on obtient directement une matrice triangulaire supérieure dont le déterminant vaut :

$$\begin{aligned}\Delta(a_1,\dots,a_n) &= \left(\prod_{i=1}^{n} a_i\right)\left(-\sum_{i=1}^{n} \frac{1}{a_i}\right) \\ &= -(a_2\cdots a_n + a_1 a_3 \cdots a_n + \cdots + a_1\cdots a_{n-1})\,.\end{aligned}$$

Cette dernière égalité reste vraie même lorsque l'un des a_i est nul : en effet, la formule $\det A = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$ montre que le déterminant d'une matrice est une fonction polynomiale, donc continue, de ses coefficients. On peut donc prolonger l'égalité obtenue en faisant tendre a_i vers 0.

✓ *Dans le cas où l'un des a_i est nul, on pouvait également faire un calcul direct en développant le déterminant selon la colonne correspondante.*

✓ Il est important de bien comprendre l'argument de continuité utilisé dans l'exemple précédent, car il permet d'éviter de refaire les calculs dans des cas particuliers. Nous utiliserons assez souvent cette propriété par la suite.

Exemple

Calculer le déterminant d'ordre $n+1$: $\begin{vmatrix} a_0 & -1 & 0 & \dots & 0 & 0 \\ a_1 & x & -1 & \dots & 0 & 0 \\ a_2 & 0 & x & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1} & 0 & 0 & \dots & x & -1 \\ a_n & 0 & 0 & \dots & 0 & x \end{vmatrix}$.

La « bonne » combinaison linéaire est ici un peu difficile à trouver.
En numérotant les lignes de 0 à n, on effectue l'opération :

$$L_n \leftarrow L_n + xL_{n-1} + x^2 L_{n-2} + \cdots + x^{n-1} L_1 + x^n L_0\,,$$

puis on développe le déterminant obtenu selon la dernière ligne (attention aux signes).

Le déterminant cherché vaut : $a_0 x^n + a_1 x^{n-1} + \dots + a_n$.

SF3.35 Développer selon une rangée

On développe un déterminant selon une ligne ou une colonne si cette rangée comporte « beaucoup » de termes nuls, ou si les cofacteurs correspondants sont faciles à calculer. Cette méthode peut être appliquée après avoir déjà transformé la matrice par des opérations élémentaires.

Exemple

Calculer le déterminant d'ordre $n+1$: $\begin{vmatrix} a_0 & -1 & 0 & \dots & 0 & 0 \\ a_1 & x & -1 & \dots & 0 & 0 \\ a_2 & 0 & x & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1} & 0 & 0 & \dots & x & -1 \\ a_n & 0 & 0 & \dots & 0 & x \end{vmatrix}$.

On peut développer directement selon la première colonne. Le déterminant sera donc égal à :

$$\sum_{i=0}^{n}(-1)^i a_i \Delta_i \,,$$

où $\Delta_0 = x^n$, $\Delta_n = (-1)^n$ *et pour* $i \in [\![1\,;n-1]\!]$ *:*

$$\Delta_i = \begin{vmatrix} -1 & 0 & & & & & & & \\ x & -1 & 0 & & & & & 0 & \\ & \ddots & \ddots & \ddots & & & & & \\ & & \ddots & \ddots & \ddots & & & & \\ & & & x & -1 & 0 & & & \\ & & & & 0 & x & -1 & & \\ & & & & & \ddots & \ddots & \ddots & \\ & & 0 & & & & \ddots & \ddots & -1 \\ & & & & & & & 0 & x \end{vmatrix} \longleftarrow i^e \text{ ligne}.$$

En calculant ce déterminant par blocs (le bloc supérieur gauche est une matrice triangulaire inférieure et le bloc inférieur droit est une matrice triangulaire supérieure), on obtient $\Delta_i = (-1)^i x^{n-i}$*, donc le déterminant cherché vaut* $\sum_{i=0}^{n} a_i x^{n-i}$.

On retrouve ainsi le résultat obtenu page 278. La matrice dont nous avons calculé le déterminant s'appelle une matrice compagnon.

SF3.36 Repérer une combinaison linéaire de rangées particulière

Si une combinaison linéaire de rangées fait apparaître une valeur particulière sur chaque rangée, on pourra faire cette combinaison linéaire et mettre la quantité obtenue en facteur. C'est par exemple le cas lorsque la somme des coefficients de chaque ligne (ou de chaque colonne) est identique.

Exemple

Calculer le déterminant d'ordre n : $D_n(a,b) = \begin{vmatrix} a & b & \dots & \dots & b \\ b & a & b & \dots & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & b & \ddots & a & b \\ b & b & b & \dots & a \end{vmatrix}$.

On remarque ici que la somme des coefficients de chaque ligne est égale à $a+(n-1)b$. Donc en faisant l'opération $C_1 \leftarrow C_1 + C_2 + \cdots + C_n$ puis en utilisant la linéarité du déterminant par rapport à la 1re colonne, on obtient :

$$D_n(a,b) = \big(a+(n-1)b\big) \begin{vmatrix} 1 & b & \dots & \dots & b \\ 1 & a & b & \dots & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & b & \ddots & a & b \\ 1 & b & b & \dots & a \end{vmatrix}.$$

En soustrayant alors la première ligne à toutes les autres, on obtient :

$$D_n(a,b) = \big(a+(n-1)b\big) \begin{vmatrix} 1 & b & \dots & \dots & b \\ 0 & a-b & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & a-b & 0 \\ 0 & 0 & 0 & \dots & a-b \end{vmatrix},$$

et enfin, en développant selon la première colonne :

$$D_n(a,b) = \big(a+(n-1)b\big)(a-b)^{n-1}.$$

SF3.37 Utiliser la multilinéarité

Il s'agit ici d'utiliser le caractère n-linéaire du déterminant, afin d'écrire un déterminant comme somme de déterminants plus simples à calculer. Cette méthode est utilisée lorsqu'un même vecteur colonne peut être mis en évidence dans toutes les colonnes de la matrice.

Exemple

Calculer le déterminant d'ordre n : $D_n(a,b) = \begin{vmatrix} a & b & \dots & \dots & b \\ b & a & b & \dots & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & b & \ddots & a & b \\ b & b & b & \dots & a \end{vmatrix}$.

Soit $V = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ et E_i les vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$. Alors

$$D_n(a,b) = \det\big((a-b)E_1 + bV, (a-b)E_2 + bV, \dots, (a-b)E_n + bV\big).$$

En développant par multilinéarité on trouve :

$$\begin{aligned} D_n(a,b) &= \det\big((a-b)E_1,(a-b)E_2,\dots,(a-b)E_n\big) \\ &\quad + \det\big(bV,(a-b)E_2,\dots,(a-b)E_n\big) + \cdots \\ &\quad + \det\big((a-b)E_1,\dots,(a-b)E_{n-1},bV\big)\,, \end{aligned}$$

les déterminants non écrits étant nuls car ils comportent au moins deux fois la même colonne bV.
En développant ensuite chacun des n derniers déterminants selon la i^e ligne (ou bien en écrivant $bV = \sum_{j=1}^{n} bE_j$ puis en utilisant la multilinéarité, comme dans l'exemple précédent), on obtient :

$$D_n(a,b) = (a-b)^n + nb(a-b)^{n-1} = (a-b)^{n-1}\big(a+(n-1)b\big)\,.$$

SF3.38 Utiliser une récurrence

Cette méthode consiste à développer un déterminant d'ordre n selon une rangée, afin de l'exprimer à l'aide du même déterminant à l'ordre $n-1$. Parfois il faut réitérer l'opération pour faire aussi apparaître le déterminant à l'ordre $n-2$.

✓ Cela exige de connaître parfaitement les résultats du cours sur les suites récurrentes linéaires d'ordre 1 et 2, rappelés en **[S8.5]** et en **[S8.6]**.

Exemple

Calculer le déterminant d'ordre n :

$$D_n(a,b) = \begin{vmatrix} a+b & ab & 0 & \dots & \dots & 0 \\ 1 & a+b & ab & 0 & \dots & 0 \\ 0 & 1 & a+b & ab & \dots & 0 \\ 0 & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \vdots & \vdots & 1 & a+b & ab \\ 0 & 0 & \dots & 0 & 1 & a+b \end{vmatrix}.$$

C'est une matrice tridiagonale, et la méthode est toujours la même : on développe selon la première colonne et on trouve : $D_n = (a+b)D_{n-1} - abD_{n-2}$ pour $n \geqslant 3$. Puisque $D_2 = (a+b)^2 - ab$ et $D_1 = a+b$, on peut étendre cette relation de récurrence à $n \geqslant 2$ en posant $D_0 = 1$.
L'équation caractéristique de la récurrence est $r^2 - (a+b)r + ab = 0$, de racines a et b.
Si $a \neq b$, on a $D_n = \lambda a^n + \mu b^n$. Puisque $D_0 = 1$ et $D_1 = a+b$, on trouve :

$$D_n(a,b) = \frac{a^{n+1} - b^{n+1}}{a-b} \quad (*)$$

Cela s'écrit aussi $D_n = \sum_{k=0}^{n} a^k b^{n-k}$, pour $b \neq a$ a priori. Mais D_n étant un polynôme en a, cette formule reste vraie pour $a = b$, d'où $D_n(a,a) = (n+1)a^n$ (on pouvait aussi passer à la limite quand $b \to a$ dans l'expression $(*)$ en utilisant la continuité de la fonction D_n).

Exemple

Calculer le déterminant d'ordre $n+1$: $\begin{vmatrix} a_0 & -1 & 0 & \dots & 0 & 0 \\ a_1 & x & -1 & \dots & 0 & 0 \\ a_2 & 0 & x & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1} & 0 & 0 & \dots & x & -1 \\ a_n & 0 & 0 & \dots & 0 & x \end{vmatrix}$.

Notons $C(a_0,a_1,\dots,a_n)$ ce déterminant.
En développant selon la dernière colonne, on obtient la relation :

$$C(a_0,a_1,\dots,a_n) = xC(a_0,a_1,\dots,a_{n-1}) + \begin{vmatrix} a_0 & -1 & 0 & \dots & 0 \\ a_1 & x & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n-2} & 0 & \dots & x & -1 \\ a_n & 0 & \dots & 0 & 0 \end{vmatrix},$$

puis en développant le déterminant d'ordre n obtenu par rapport à la dernière ligne, on obtient :

$$\begin{aligned} C(a_0,a_1,\dots,a_n) &= xC(a_0,a_1,\dots,a_{n-1}) + (-1)^{n+1}a_n(-1)^{n-1} \\ &= xC(a_0,a_1,\dots,a_{n-1}) + a_n\,. \end{aligned}$$

Il est alors facile de montrer par récurrence sur n que l'on a :

$$C(a_0,a_1,\dots,a_n) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n\,,$$

résultat que nous avons déjà obtenu par d'autres méthodes (à vous de choisir celle qui vous plaît le plus !)

SF3.39 Utiliser le caractère polynomial du déterminant

Les formules de développement du déterminant selon une rangée montrent que le déterminant d'une matrice est la somme de produits des coefficients de cette matrice ; c'est donc une fonction polynomiale des coefficients. Nous avons déjà utilisé cette propriété dans certains exemples précédents (pour prolonger la valeur du déterminant en un point, par continuité) ; voici d'autres exemples de son utilité.

Exemple

Calculer $\Delta = \begin{vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{vmatrix}$.

En développant (fictivement) ce déterminant par rapport à la première colonne, on s'aperçoit qu'il s'agit d'une fonction polynomiale en a, de degré 4 et unitaire (car le terme en a^4 s'obtient en faisant le produit des éléments diagonaux).

En faisant l'opération $C_1 \leftarrow C_1+C_2+C_3+C_4$, puis en utilisant la linéarité du déterminant par rapport à cette colonne, on obtient que ce polynôme en a est divisible par $a+b+c+d$.

De même, en considérant les opérations $C_1 \leftarrow C_1 + C_2 - C_3 - C_4$, $C_1 \leftarrow C_1 - C_2 + C_3 - C_4$ et $C_1 \leftarrow C_1 - C_2 - C_3 + C_4$, on voit qu'il est aussi divisible par $a+b-c-d$, $a-b+c-d$ et $a-b-b+d$.

Finalement, $\Delta = (a+b+c+d)(a+b-c-d)(a-b+c-d)(a-b-c+d)$.

Exemple

Soit $x \in \mathbb{C}$.

Calculer le déterminant d'ordre $n+1$: $D_n(x) = \begin{vmatrix} 1 & & & (1) \\ & 1-x & & \\ & & \ddots & \\ (1) & & & n-x \end{vmatrix}$.

C'est un polynôme de degré n en x, qui s'annule pour $x = 0, \dots, n-1$ (car il y a alors une ligne égale à la première).
Par ailleurs, le terme en x^n est obtenu en faisant le produit des éléments diagonaux (cela peut se justifier rigoureusement en revenant à la formule du cours : $\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$), donc le coefficient dominant de ce polynôme est égal à $(-1)^n$.

En conclusion : $D_n(x) = (-1)^n \prod_{k=0}^{n-1} (x-k)$.

SF3.40 Utiliser la dérivation du déterminant

Nous avons vu en **[S10.2]** que si les coefficients d'une matrice $M(t)$ sont des fonctions dérivables de t, alors l'application $t \mapsto \det M(t)$ est dérivable et que sa dérivée vaut, en notant $C_j(t)$ les colonnes de $M(t)$:

$$\sum_{j=1}^{n} \det\big(C_1(t), \dots, C_{j-1}(t), C_j'(t), C_{j+1}(t), \dots, C_n(t)\big),$$

Exemple

Pour tout réel x, on pose : $D_n(x) = \begin{vmatrix} x & 1 & & & 0 \\ \frac{x^2}{2!} & x & 1 & & \\ \frac{x^3}{3!} & \frac{x^2}{2!} & x & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ \frac{x^n}{n!} & \cdots & \cdots & \frac{x^2}{2!} & x \end{vmatrix}$.

Montrer que D_n est dérivable, calculer $D_n'(x)$ et en déduire l'expression de $D_n(x)$.

Les coefficients de la matrice donnée étant des fonctions dérivables de x, il en est de même de D_n.

On utilise la formule rappelée ci-dessus ; pour tout $j \in [\![1\,;n-1]\!]$, on a $C_j'(x) = C_{j+1}(x)$ donc tous les déterminants dans la somme sont nuls sauf le dernier, qui est égal à $\det\big(C_1(t), \dots, C_{n-1}(t), C_n'(t)\big)$. Ainsi :

$$D_n'(x) = \begin{vmatrix} x & 1 & & & 0 \\ \frac{x^2}{2!} & x & 1 & & \\ \frac{x^3}{3!} & \frac{x^2}{2!} & x & \ddots & \\ \vdots & & \ddots & \ddots & 0 \\ \frac{x^n}{n!} & \cdots & \cdots & \frac{x^2}{2!} & 1 \end{vmatrix} = D_{n-1}(x)\,,$$

après avoir développé selon la dernière colonne.
On en déduit facilement par récurrence $D_n(x) = \dfrac{x^n}{n!}$.

SF3.41 Faire des calculs par blocs

On connaît bien sûr la formule du déterminant par blocs ; les exemples ci-dessous montrent qu'il est aussi possible de faire des opérations élémentaires par blocs (en le justifiant rigoureusement).

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $(a,b,c,d) \in \mathbb{K}^4$.

Calculer le déterminant de la matrice d'ordre $2n$: $\begin{bmatrix} aA & bA \\ cA & dA \end{bmatrix}$.

Supposons d'abord $d \neq 0$. *Les opérations* $C_i \longleftarrow C_i - \dfrac{c}{d}C_{i+n}$ *pour* $1 \leqslant i \leqslant n$ *donnent :*

$$\begin{vmatrix} aA & bA \\ cA & dA \end{vmatrix} = \begin{vmatrix} \left(a - \frac{bc}{d}\right) A & bA \\ 0 & dA \end{vmatrix} = \det\left(\left(a - \frac{bc}{d}\right) A\right) \times \det(dA)$$
$$= \left(a - \frac{bc}{d}\right)^n d^n (\det A)^2 = (ad - bc)^n (\det A)^2\,,$$

et l'égalité reste vraie pour $d = 0$ *puisqu'il s'agit de deux polynômes en* d.

Exemple

Soient A, $B \in \mathcal{M}_n(\mathbb{R})$. On pose $\Delta = \begin{vmatrix} A & B \\ -B & A \end{vmatrix}$. Montrer que $\Delta \geqslant 0$.

On fait les calculs dans $\mathbb{C}$. *Les opérations* $C_j \longleftarrow C_j + \mathrm{i}C_{j+n}$ *pour* $1 \leqslant j \leqslant n$ *donnent :*

$$\begin{vmatrix} A & B \\ -B & A \end{vmatrix} = \begin{vmatrix} A + \mathrm{i}B & B \\ \mathrm{i}A - B & A \end{vmatrix},$$

puis les opérations $L_{k+n} \longleftarrow L_{k+n} - \mathrm{i}L_k$ *pour* $1 \leqslant k \leqslant n$ *donnent :*

$$\begin{vmatrix} A + \mathrm{i}B & B \\ \mathrm{i}A - B & A \end{vmatrix} = \begin{vmatrix} A + \mathrm{i}B & B \\ 0 & A - \mathrm{i}B \end{vmatrix}.$$

Donc $\Delta = \det(A + \mathrm{i}B)\det(A - \mathrm{i}B)$.
Or pour toute matrice M *carrée à coefficients complexes, on a* $\det(\overline{M}) = \overline{\det M}$ *: cela se justifie aisément en utilisant la formule* $\det M = \displaystyle\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{k=1}^{n} m_{\sigma(k)k}$ *et les propriétés de la conjugaison dans* $\mathbb{C}$.

Donc A et B étant réelles, $\det(A-\mathrm{i}B) = \overline{\det(A+\mathrm{i}B)}$ *et* $\Delta = |\det(A+\mathrm{i}B)|^2$ *est bien positif.*

Exemple

Soient $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$. En faisant d'astucieux calculs par blocs, comparer les polynômes caractéristiques de AB et de BA.

Posons $U = \begin{bmatrix} x\mathrm{I}_n & A \\ B & \mathrm{I}_p \end{bmatrix} \in \mathcal{M}_{n+p}(\mathbb{K})$. *Puisque :*

$$\begin{bmatrix} x\mathrm{I}_n & A \\ B & \mathrm{I}_p \end{bmatrix} \times \begin{bmatrix} \mathrm{I}_n & 0 \\ -B & \mathrm{I}_p \end{bmatrix} = \begin{bmatrix} x\mathrm{I}_n - AB & A \\ 0 & \mathrm{I}_p \end{bmatrix},$$

on a : $\det U = \chi_{AB}(x)$. *D'autre part, l'égalité :*

$$\begin{bmatrix} x\mathrm{I}_n & A \\ B & \mathrm{I}_p \end{bmatrix} \times \begin{bmatrix} \mathrm{I}_n & -A \\ 0 & x\mathrm{I}_p \end{bmatrix} = \begin{bmatrix} x\mathrm{I}_n & 0 \\ B & x\mathrm{I}_p - BA \end{bmatrix}$$

donne : $x^p \det U = x^n \chi_{BA}(x)$.

On en déduit : $\chi_{AB} = X^{n-p}\chi_{BA}$.

Cela permet de montrer que les matrices AB et BA ont les mêmes valeurs propres non nulles avec le même ordre de multiplicité, et aussi que, lorsque $p = n$, $\chi_{AB} = \chi_{BA}$.

À vous de jouer !

Exercice 3.1 **SF3.2**

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que pour toutes matrices M, $N \in \mathcal{M}_n(\mathbb{K})$ on ait $\mathrm{tr}(AMN) = \mathrm{tr}(MAN)$.

Montrer que A est une matrice scalaire.

Exercice 3.2 **SF3.27**

Dans $\mathbb{R}^n$, soit $u = (u_1, u_2, \ldots, u_n)$ tel que $\sum_{k=1}^{n} u_k^2 = 1$, D la droite engendrée par u et H l'hyperplan d'équation $u_1x_1 + u_2x_2 + \ldots + u_nx_n = 0$.

Déterminer les matrices dans la base canonique de la projection sur H parallèlement à D, puis de la symétrie par rapport à D parallèlement à H.

Exercice 3.3 **SF3.22**

Soit $C \in \mathcal{M}_{n,1}(\mathbb{K})$ et $L \in \mathcal{M}_{1,n}(\mathbb{K})$.

Déterminer une condition nécessaire et suffisante pour que la matrice $A = \mathrm{I}_n + CL$ soit inversible et calculer alors son inverse.

Exercice 3.4 **SF3.3** **SF3.14**

Soit $A \in \mathcal{M}_n(\mathbb{K})$, de rang r. On note $F = \{B \in \mathcal{M}_n(\mathbb{K}) \mid ABA = 0_n\}$.
Montrer que F est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ et en déterminer la dimension.

Exercice 3.5 **SF3.10** **SF3.13** **SF3.14**

Soit $M \in \mathcal{M}_{p+q}(\mathbb{K})$ une matrice partitionnée par blocs : $M = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$, avec $A \in \mathcal{M}_p(\mathbb{K})$, $B \in \mathcal{M}_{q,p}(\mathbb{K})$ et $C \in \mathcal{M}_q(\mathbb{K})$.

1) Montrer que : $\operatorname{rg}(M) \geqslant \operatorname{rg}(A) + \operatorname{rg}(C)$.

2) Montrer que, si A est inversible, il y a égalité.

Exercice 3.6 **SF3.26**

Soit E un $\mathbb{K}$-espace vectoriel de dimension $3n$ ($n \in \mathbb{N}^*$), et u un endomorphisme de E tel que : $\operatorname{rg} u = 2n$ et $u^3 = 0$.

1) Montrer que $\operatorname{Ker} u = \operatorname{Im} u^2$ et que $\operatorname{Ker} u^2 = \operatorname{Im} u$.

2) Montrer qu'il existe une base $\mathscr{B}$ de E où la matrice de u s'écrit par blocs :

$$M_{\mathscr{B}}(u) = \begin{bmatrix} 0_n & 0_n & 0_n \\ \mathrm{I}_n & 0_n & 0_n \\ 0_n & \mathrm{I}_n & 0_n \end{bmatrix}.$$

Exercice 3.7 **SF3.32**

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Déterminer toutes les matrices $X \in \mathcal{M}_n(\mathbb{R})$ telles que :

$$X + {}^{\mathrm{t}}X = \operatorname{tr}(X)A\,.$$

Exercice 3.8 **SF3.26**

Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\varphi : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ M & \longmapsto & AM + MA\,. \end{cases}$

Montrer que φ est un endomorphisme de $\mathcal{M}_n(\mathbb{K})$, et calculer sa trace en fonction de celle de A.

Exercice 3.9 **SF3.20** **SF3.39**

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{C})$, telle que : $\forall\, i \in [\![1\,;n]\!]$, $\displaystyle\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| < |a_{ii}|$ (une telle matrice est dite *à diagonale strictement dominante*).

1) Montrer que A est inversible (*on pourra raisonner par l'absurde, en supposant qu'il existe $X \neq 0$ tel que $AX = 0$*).

2) On suppose de plus que A est à coefficients réels, et que ses éléments diagonaux sont strictement positifs. Montrer que $\det A > 0$.

Exercice 3.10 — SF3.1

Soit $T = (t_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$ définie par : $t_{ij} = \begin{cases} 1 & \text{si } i \text{ divise } j \\ 0 & \text{sinon} \end{cases}$.

On note φ l'indicateur d'Euler, et D la matrice diagonale d'ordre n :

$$D = \operatorname{diag}\big(\varphi(1), \ldots, \varphi(n)\big).$$

On rappelle également la propriété suivante de l'indicateur d'Euler :

$$\forall\, n \in \mathbb{N}^*,\ n = \sum_{d \mid n} \varphi(d)\,.$$

1) Calculer le coefficient d'indice (i,j) de la matrice ${}^{\mathrm{t}}TDT$ en fonction de $i \wedge j$.

2) En déduire le déterminant de la matrice $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ définie par :

$$\forall\, (i,j) \in [\![1\,;n]\!]^2,\ a_{ij} = i \wedge j\,.$$

Exercice 3.11 — SF3.13 | SF3.35 | SF3.38

Soient $n \in \mathbb{N}^*$, $z \in \mathbb{C}$ et $A_n = \begin{pmatrix} 1+z^2 & z & 0 & \ldots & 0 \\ z & 1+z^2 & z & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & z & 1+z^2 & z \\ 0 & \ldots & 0 & z & 1+z^2 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$.

Déterminer le rang et le déterminant de A_n.

Exercice 3.12 — SF3.37

Calculer le déterminant d'ordre n : $\Delta = \begin{vmatrix} a_1 + x & a_1 & \ldots & a_1 \\ a_2 & a_2 + x & \ldots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_n & \ldots & a_n + x \end{vmatrix}$.

Thème 4 - Réduction des endomorphismes

■ Comment déterminer les éléments propres d'un endomorphisme ?

Les méthodes pour déterminer les valeurs propres et vecteurs propres d'un endomorphisme sont assez nombreuses, et il n'est pas toujours facile de savoir laquelle est la mieux adaptée à un problème précis. Par exemple, dans certains cas, le calcul du polynôme caractéristique est incontournable, alors que dans d'autres cas il conduit à des calculs inextricables.
Il est donc important de traiter de nombreux exemples différents afin de pouvoir reconnaître des situations types.
Lorsque l'espace vectoriel n'est pas de dimension finie, le nombre de méthodes est très réduit puisque toutes les notions basées sur le calcul matriciel (polynôme caractéristique) et la dimension des sous-espaces propres n'ont alors pas de sens. D'une certaine façon, cela simplifie la tâche !

N'oubliez pas que, lorsque vous utilisez un polynôme (caractéristique ou annulateur) pour chercher les valeurs propres, il faut faire attention au corps de base : les résultats diffèrent selon qu'il s'agit de $\mathbb{R}$ ou de $\mathbb{C}$.

SF4.1 Revenir à la définition

Il s'agit donc, pour un endomorphisme u d'un espace vectoriel E, de résoudre l'équation $u(x) = \lambda x$, d'inconnues $\lambda \in \mathbb{K}$ et $x \in E \setminus \{0\}$. Plus précisément, on cherche les valeurs du paramètre λ pour lesquelles cette équation admet une solution x non nulle.
Cette méthode permet donc d'obtenir simultanément les valeurs propres et les vecteurs propres de u. Il s'agit pratiquement de la seule méthode utilisable lorsque E n'est pas de dimension finie. Mais elle peut être très intéressante également lorsque E est de dimension finie car elle évite de rechercher le polynôme caractéristique.

Exemple

Soit $A \neq 0_n \in \mathcal{M}_n(\mathbb{K})$ donnée, et soit u l'endomorphisme de $\mathcal{M}_n(\mathbb{K})$ défini par :

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}), \quad u(M) = \mathrm{tr}(A)M - \mathrm{tr}(M)A\,.$$

Déterminer ses éléments propres. u est-il diagonalisable ?

Il s'agit donc de trouver les scalaires $\lambda \in \mathbb{K}$ *pour lesquels il existe une matrice* $M \neq 0_n$ *telle que :*

$$\mathrm{tr}(A)M - (\mathrm{tr}\, M)A = \lambda M \quad (*)$$

En appliquant la fonction trace, *qui est linéaire, à la relation* $(*)$ *on obtient :*

$$\mathrm{tr}(A)\,\mathrm{tr}(M) - \mathrm{tr}(M)\,\mathrm{tr}(A) = \lambda\,\mathrm{tr}(M)$$

d'où nécessairement $\lambda\,\mathrm{tr}(M) = 0$. *Deux cas se présentent donc.*

- *Si* $\lambda \neq 0$ *alors* $\mathrm{tr}(M) = 0$ *et la relation* $(*)$ *donne alors* $\mathrm{tr}(A)M = \lambda M$ *d'où* $\lambda = \mathrm{tr}(A)$ *puisque* $M \neq 0$.
  Réciproquement, si $\lambda = \mathrm{tr}(A)$, *toute matrice* M *de trace nulle vérifie* $(*)$.

Ainsi, $tr(A)$ est valeur propre de u, et le sous-espace propre associé est l'ensemble des matrices de trace nulle ; comme c'est le noyau d'une forme linéaire non nulle, c'est un hyperplan de $\mathcal{M}_n(\mathbb{K})$.

- *Pour $\lambda = 0$, la relation $(*)$ équivaut à $\mathrm{tr}(A)M = \mathrm{tr}(M)A$. Deux cas se présentent donc à nouveau.*
  - *Si $\mathrm{tr}(A) = 0$, alors on retrouve l'égalité $\mathrm{tr}(M) = 0$. Dans ce cas, 0 est la seule valeur propre de u, et u n'est pas diagonalisable (sinon, u serait nulle, ce qui est exclu puisque $A \neq 0$).*
  - *Si $\mathrm{tr}(A) \neq 0$ on a $M = \dfrac{\mathrm{tr}(M)}{\mathrm{tr}(A)}A$ donc il existe $k \in \mathbb{K}$ tel que $M = kA$; réciproquement, si $M = kA$, on vérifie aisément que $u(M) = 0$ donc dans ce cas 0 est valeur propre de u, et le sous-espace propre associé est la droite vectorielle engendrée par A.*

Conclusion : u est diagonalisable si et seulement si $\mathrm{tr}(A) \neq 0$ (en effet, dans ce cas, nous avons trouvé deux sous-espaces propres dont la somme des dimensions est égale à celle de $\mathcal{M}_n(\mathbb{K})$).

Exemple

On note $E = \mathscr{C}([0\,;1],\mathbb{R})$, espace vectoriel des fonctions continues sur $[0\,;1]$.
Pour tout $f \in E$, on définit la fonction $F : [0\,;1] \longrightarrow \mathbb{R}$ par :

$$F(x) = \begin{cases} \dfrac{1}{x}\displaystyle\int_0^x f(t)\,\mathrm{d}t & \text{si } x \in]0\,;1] \\ f(0) & \text{si } x = 0\,. \end{cases}$$

Montrer que $F \in E$ et que l'application $\varphi\colon f \longmapsto F$ est un endomorphisme de E.
Déterminer les valeurs propres et les vecteurs propres de φ.

- *f étant continue, l'application $G : x \mapsto \displaystyle\int_0^x f(t)\,\mathrm{d}t$ est la primitive de f s'annulant en 0. Donc $F : x \mapsto \dfrac{1}{x}G(x)$ est de classe $\mathscr{C}^1$ sur $]0\,;1]$.*
  De plus,

$$\lim_{x\to 0^+} F(x) = \lim_{x\to 0^+} \frac{G(x) - G(0)}{x - 0} = G'(0) = f(0) = F(0)\,,$$

  donc F est continue en 0^+ et, par suite, sur $[0\,;1]$.
  On a donc bien $F \in E$. La linéarité de φ découle directement de la linéarité de l'intégration, donc φ est bien un endomorphisme de E.
- *Soit f un vecteur propre de φ, alors $F = \lambda f$ donc pour tout $x \in]0\,;1]$ on a : $\displaystyle\int_0^x f = \lambda x f(x)$ $(*)$, et cette égalité reste vraie pour $x = 0$.*
  - *Pour $\lambda = 0$ on a $\displaystyle\int_0^x f = 0$ pour tout x d'où, en dérivant, $f = 0$. Ainsi, 0 n'est pas une valeur propre.*
  - *Lorsque $\lambda \neq 0$, la relation $f(x) = \dfrac{1}{\lambda x}\displaystyle\int_0^x f$ pour $x \in]0\,;1]$ implique que f est dérivable sur $]0\,;1]$; on a alors en dérivant l'égalité $(*)$, $f(x) = \lambda f(x) + \lambda x f'(x)$ pour $x \in]0\,;1]$.*

f est donc solution sur $]0\,;1]$ de l'équation différentielle :

$$f'(x) = \frac{1-\lambda}{\lambda x} f(x)\,.$$

Les solutions sont les fonctions de la forme $x \mapsto Cx^{\frac{1-\lambda}{\lambda}}$. Il faut de plus que cette solution se prolonge par continuité en 0, ce qui exige $\dfrac{1-\lambda}{\lambda} \geqslant 0$, soit $\lambda \in]0\,;1]$.

Réciproquement, un calcul simple d'intégrale montre que, si $\lambda \in]0\,;1]$ et si $f(x) = Cx^{\frac{1-\lambda}{\lambda}}$ pour tout $x \in [0\,;1]$, alors $f \in E$ et $F = \lambda f$.

En conclusion, $\mathrm{Sp}(\varphi) =]0\,;1]$ *et* $\forall\, \lambda \in]0\,;1]$, $E_\lambda(\varphi) = \mathrm{Vect}\left(x \mapsto x^{\frac{1-\lambda}{\lambda}}\right)$.

SF4.2 Utiliser un polynôme annulateur

On sait que si u est un endomorphisme d'un espace vectoriel E (de dimension finie ou non) et si P est un polynôme tel que $P(u) = 0_{\mathscr{L}(E)}$, alors *nécessairement*, pour toute valeur propre λ de u on a $P(\lambda) = 0$.
Si l'on connaît un tel polynôme annulateur P, on en déterminera alors les racines λ puis, pour chaque valeur trouvée, on résoudra l'équation $u(x) = \lambda x$. Si cette équation admet des solutions $x \neq 0_E$, alors λ sera effectivement une valeur propre de u, et on aura aussi trouvé les sous-espaces propres correspondants.

Exemple

Soit a un réel non nul et $A = \begin{pmatrix} 1 & a & a^2 \\ \frac{1}{a} & 1 & a \\ \frac{1}{a^2} & \frac{1}{a} & 1 \end{pmatrix}$. Calculer A^2 et en déduire les valeurs propres de A.

On vérifie que $A^2 = A + 2\mathrm{I}_3$.
Le polynôme $P = X^2 - X - 2$ est donc annulateur de A. Or $P = (X+1)(X-2)$, donc P est scindé à racines simples dans $\mathbb{R}$. La matrice A est donc diagonalisable et $\mathrm{Sp}(A) \subset \{-1,2\}$.
Si l'une des racines de P n'était pas valeur propre de A, A serait diagonalisable avec une seule valeur propre et serait donc semblable donc égale à une matrice scalaire, ce qui n'est pas le cas.
Ainsi, $\mathrm{Sp}(A)$ *est exactement l'ensemble* $\{-1,2\}$.

✓ Il est important de retenir la propriété utilisée ici : si $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable et ne possède qu'une seule valeur propre λ elle est nécessairement égale à $\lambda \mathrm{I}_n$. En effet, étant diagonalisable elle est semblable à $\lambda \mathrm{I}_n$ donc il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A = P(\lambda \mathrm{I}_n)P^{-1}$ d'où $A = \lambda \mathrm{I}_n$.

Exemple

Soit $A \neq 0_n \in \mathcal{M}_n(\mathbb{K})$ donnée, et soit u l'endomorphisme de $\mathcal{M}_n(\mathbb{K})$ défini par :

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}), \quad u(M) = \mathrm{tr}(A)M - \mathrm{tr}(M)A\,.$$

Déterminer ses éléments propres. u est-il diagonalisable ?

Cet exemple a déjà été traité auparavant, mais vous allez pouvoir constater que l'utilisation d'un polynôme annulateur simplifie notablement le problème.
Pour trouver un polynôme annulateur de u, on calcule :

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}),\ u^2(M) = \mathrm{tr}(A)u(M) - \mathrm{tr}(M)\underbrace{u(A)}_{=0} = \mathrm{tr}(A)u(M)\,,$$

donc $u^2 = \mathrm{tr}(A)u$ et le polynôme $X(X - \mathrm{tr}(A))$ est annulateur de u.

- *Si $\mathrm{tr}(A) \neq 0$, ce polynôme est scindé à racines simples, donc u est diagonalisable, et comme ce n'est pas une homothétie, on a $\mathrm{Sp}(u) = \{0, \mathrm{tr}(A)\}$. Il est facile ensuite de vérifier que le sous-espace propre associé à la valeur propre $\mathrm{tr}(A)$ est $\mathrm{Ker}\, u = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \mathrm{tr}(M) = 0\}$; c'est donc un hyperplan de $\mathcal{M}_n(\mathbb{K})$.*
  Le second sous-espace propre, celui associé à la valeur propre 0, ne peut donc qu'être une droite, et puisque l'on a vu que $u(A) = 0$, il s'agit forcément de la droite $\mathbb{K}.A$.
- *Si $\mathrm{tr}(A) = 0$, u possède une seule valeur propre qui est 0. Comme u n'est pas l'endomorphisme nul, il n'est pas diagonalisable.*

✓ La difficulté dans cet exemple était surtout de *penser* à utiliser un polynôme annulateur ; cela n'est pas toujours indiqué dans l'énoncé.

SF4.3 Utiliser le polynôme caractéristique

Beaucoup considèrent que pour étudier les éléments propres d'un endomorphisme d'un espace vectoriel de dimension finie, ou d'une matrice, il faut systématiquement calculer le polynôme caractéristique. Les quelques exemples précédents montrent qu'il n'en est rien. En effet, cette méthode présente deux inconvénients majeurs : il faut calculer un déterminant, ce qui peut être assez difficile, et il faut ensuite savoir trouver les racines du polynôme caractéristique.

Cette méthode est donc à utiliser avec prudence. Elle est cependant recommandée lorsque la dimension de l'espace (ou la taille de la matrice) est petite (inférieure ou égale à 4), ou lorsque la matrice est d'une forme « classique ».

Cependant, même dans ces cas simples, nous avons constaté que beaucoup d'étudiants n'arrivaient pas à terminer sans erreur le calcul d'un déterminant 3×3 ou 4×4 ! Pour cela, il n'y a qu'une seule solution : s'entraîner aux calculs, en essayant de les mener de la façon la plus astucieuse possible (utiliser des combinaisons linéaires sur les rangées d'une matrice, factoriser plutôt que de développer, faire des calculs par blocs etc...). Pour cette raison, nous n'avons pas traité ici d'exemples de calcul de polynôme caractéristique de matrices de tailles 3 et 4 : c'est à vous de vous exercer, vous trouverez des exemples dans la partie « À vous de jouer ! »

✓ Il est important de se rappeler que le polynôme caractéristique d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ commence par : $\chi_A = X^n - \mathrm{tr}(A)X^{n-1} + \cdots$. Cela permet une vérification rapide de ses calculs.

Si l'on essaye en général d'éviter de calculer le polynôme caractéristique d'une matrice d'ordre n, il est des cas où ce calcul est assez simple : il ne faut donc pas hésiter à s'en servir. Il est aussi des cas où il s'agit de la seule méthode possible.

Exemple

Soient $a_0, a_1, \dots, a_{n-1}$ des nombres complexes, P le polynôme $P = X^n + \sum_{i=0}^{n-1} a_i X^i$ et C la matrice de $\mathcal{M}_n(\mathbb{C})$ définie par :

$$C = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & 0 & & \vdots & -a_1 \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ & \ddots & \ddots & \ddots & \vdots & \vdots \\ & 0 & \ddots & \ddots & 0 & -a_{n-2} \\ & & & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

1) Calculer le polynôme caractéristique de C.

2) Trouver les éléments propres de ${}^tC$ et en déduire une condition nécessaire et suffisante pour que C soit diagonalisable.

1) *Le polynôme caractéristique est :*

$$\chi_C(X) = \begin{vmatrix} X & \cdots & \cdots & \cdots & 0 & a_0 \\ -1 & X & 0 & & \vdots & a_1 \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ & \ddots & \ddots & \ddots & \vdots & \vdots \\ & 0 & \ddots & \ddots & X & a_{n-2} \\ & & & 0 & -1 & X + a_{n-1} \end{vmatrix}.$$

Nous avons déjà calculé ce déterminant de trois façons différentes dans le chapitre précédent (pages 278, 279 et 282). Nous avions trouvé : $\chi_C(X) = P(X)$ où P est le polynôme de l'énoncé.

✓ *La matrice C s'appelle une matrice* compagnon *; elle intervient dans un certain nombre de problèmes, et il est bon de savoir calculer son polynôme caractéristique.*

2) *Les valeurs propres de C, donc aussi celles de ${}^tC$, sont les racines (dans $\mathbb{C}$) du polynôme P. Soit λ une de ces racines, et $X = {}^t\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$.*

$$ {}^tCX = \lambda X \iff \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

$$\iff \begin{cases} x_2 & = \lambda x_1 \\ x_3 & = \lambda x_2 \\ \dots & = \\ x_n & = \lambda x_{n-1} \\ -a_0 x_1 - a_1 x_2 - \cdots a_{n-1} x_n & = \lambda x_n \end{cases}$$

$$ {}^tCX = \lambda X \iff \begin{cases} \forall\, i \in [\![1\,;n]\!],\ x_i = \lambda^{i-1} x_1 \\ (-a_0 - a_1\lambda - \cdots - a_{n-1}\lambda^{n-1})x_1 = \lambda^n x_1\,. \end{cases} $$

La dernière de ces équations est automatiquement vérifiée puisque $P(\lambda) = 0$*, donc le sous-espace propre de* ${}^tC$ *associé à la valeur propre* λ *est la droite vectorielle de base le vecteur* $\begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}$.

Pour que ${}^tC$ *(donc* C*) soit diagonalisable, il faut et il suffit que la somme des dimensions des sous-espaces propres soit égale à* n*, ce qui équivaut à dire que le polynôme* P *possède exactement* n *racines distinctes (il sera donc scindé à racines simples).*

Dans ce cas, si $(\lambda_1, \ldots, \lambda_n)$ *sont les racines de* P*, on vient de montrer que l'on a :* ${}^tC = QDQ^{-1}$ *avec* $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *et* $Q = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$*. C'est une matrice de Vandermonde.*

SF4.4 Utiliser la trace et le déterminant

Il s'agit ici d'utiliser la propriété suivante.

> Si u est un endomorphisme d'un espace vectoriel de dimension finie dont le polynôme caractéristique est scindé, alors $\operatorname{tr} u$ est égale à la somme des valeurs propres de u (comptées avec leurs ordres de multiplicité) et $\det u$ est égal à leur produit.

Cette propriété peut permettre d'éviter le calcul du polynôme caractéristique dans certains cas.

> ✓ Essayez de retenir les « astuces » utilisées dans les exemples qui suivent, cela peut vous faire gagner beaucoup de temps dans les calculs.

Exemple

Trouver les valeurs propres de la matrice $A = \begin{pmatrix} 6 & -6 & 5 \\ -4 & -1 & 10 \\ 7 & -6 & 4 \end{pmatrix}$.

On remarque ici que la somme des termes de chaque ligne est égale à 5*; cela signifie, matriciellement, que :*

$$ A \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, $$

donc 5 *est valeur propre de* A *et le vecteur* ${}^t\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ *en est un vecteur propre associé.*

On calcule alors $\operatorname{tr} A = 9$ *et* $\det A = -25$*. Puisque la somme des valeurs propres (dans* $\mathbb{C}$*) est égale à* $\operatorname{tr} A$ *et le produit à* $\det A$*, les deux autres valeurs propres* λ *et* μ *de* A *vérifient :*

$$ \lambda + \mu = 4 \quad \text{et} \quad \lambda\mu = -5\,, $$

d'où $\lambda = 5$ *et* $\mu = -1$.

Ainsi les valeurs propres de A *sont* -1 *(simple) et* 5 *(double).*

Sans calculer les sous-espaces propres, on peut remarquer que la matrice $A - 5\mathrm{I}_3$ *est de rang* 2 *; donc* $\dim \mathrm{Ker}(A - 5\mathrm{I}_3) = 1$*, c'est-à-dire que le sous-espace propre associé à la valeur propre* 5 *est de dimension* 1 *; il en résulte que* A *n'est pas diagonalisable.*

Exemple

Trouver les valeurs propres de la matrice $A = \begin{pmatrix} 1 & -1 & -2 \\ 0 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}$.

On remarque ici que la somme des termes de chaque colonne est égale à 2 *; cela peut s'écrire, matriciellement :*

$$ {}^{\mathrm{t}}A \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, $$

donc 2 *est valeur propre de* ${}^{\mathrm{t}}A$*. Puisque les matrices* A *et* ${}^{\mathrm{t}}A$ *ont mêmes valeurs propres, on en déduit que* 2 *est aussi valeur propre de* A*.*

On procède alors comme dans l'exemple précédent : puisque $\mathrm{tr}\, A = 6$ *et* $\det A = 8$*, on en déduit facilement que* 2 *est la seule valeur propre, d'ordre de multiplicité* 3*.*

Il en résulte que la matrice A *n'est pas diagonalisable : en effet, si elle l'était elle serait semblable à* $2\mathrm{I}_3$*, donc égale à* $2\mathrm{I}_3$*, ce qui n'est pas le cas.*

Exemple

Trouver les éléments propres de la matrice $A = \begin{pmatrix} 6 & -6 & 5 \\ 14 & -13 & 10 \\ 7 & -6 & 4 \end{pmatrix}$.

On remarque ici que $A + \mathrm{I}_3 = \begin{pmatrix} 7 & -6 & 5 \\ 14 & -12 & 10 \\ 7 & -6 & 5 \end{pmatrix}$ *donc* $\mathrm{rg}(A + \mathrm{I}_3) = 1$ *puisque les trois colonnes de* $A + \mathrm{I}_3$ *sont proportionnelles.*

Il en résulte que $\dim \mathrm{Ker}(A + \mathrm{I}_3) = 2$*; or* $\mathrm{Ker}(A + \mathrm{I}_3)$ *n'est autre que le sous-espace propre associé à la valeur propre* -1 *; donc* -1 *est valeur propre d'ordre de multiplicité supérieure ou égale à* 2*.*

Il ne reste donc plus qu'à trouver une valeur propre, et comme $\mathrm{tr}\, A = -3$*, cette dernière valeur propre est aussi égale à* -1*.*

Ainsi -1 *est valeur propre triple ; comme la dimension du sous-espace propre associé est égale à* 2*, on peut en déduire que* A *n'est pas diagonalisable.*

SF4.5 Utiliser la trace et le rang

Si un endomorphisme d'un espace vectoriel de dimension n (ou une matrice carrée d'ordre n) est de rang r, on sait que la dimension de son noyau est égale à $n - r$. Compte tenu de la relation entre l'ordre de multiplicité d'une valeur propre et la dimension du sous-espace propre correspondant **[S4.5]**, cela implique, lorsque $r < n$, que 0 est valeur propre d'ordre *au moins* $n - r$.

Il ne reste donc plus que r valeurs propres à déterminer ; lorsque r est « petit », cela peut se faire en utilisant $\operatorname{tr} A$ voire $\operatorname{tr} A^2$, à l'aide de la propriété suivante.

Si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres complexes de A, on a :

$$\forall\, k \in \mathbb{N},\ \sum_{i=1}^{n} \lambda_i^k = \operatorname{tr}(A^k)\,.$$

Dans le cas $k = 1$, ce résultat est classique (cf. **[S4.4]**).
Pour le démontrer dans le cas général, on utilise le fait que toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable : si $A = PTP^{-1}$ avec T triangulaire supérieure, alors $A^k = PT^kP^{-1}$. Et puisque les coefficients de la diagonale de T sont les λ_i, ceux de la diagonale de T^k sont les λ_i^k d'où $\operatorname{tr}(A^k) = \operatorname{tr}(T^k) = \sum_{i=1}^{n} \lambda_i^k$.

Exemple

Trouver les éléments propres de la matrice $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{pmatrix}$.

Les 4 colonnes de la matrice étant égales, le rang de A est égal à 1. On en déduit, d'après le théorème du rang, que $\operatorname{Ker} A$ est de dimension 3 donc 0 est valeur propre de A d'ordre $\geqslant 3$. Mais puisque la somme des 4 valeurs propres de A (dans $\mathbb{C}$) est égale à la trace de A c'est-à-dire à 10, on en déduit que les valeurs propres de A sont 0 (triple) et 10 (simple).

Si X est le vecteur colonne $X = {}^{\mathrm{t}}\begin{pmatrix} x & y & z & t \end{pmatrix}$, le système $AX = 0$ équivaut à $x + y + z + t = 0$: c'est l'équation d'un hyperplan.

On remarque alors que $A \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \\ 30 \\ 40 \end{pmatrix} = 10 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ donc le sous-espace propre associé à la valeur propre 10 (qui est nécessairement une droite puisque cette valeur propre est simple) est la droite vectorielle engendrée par le vecteur colonne ${}^{\mathrm{t}}\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$.

Enfin, les deux sous-espaces propres trouvés étant supplémentaires (la droite n'est pas incluse dans l'hyperplan), on peut dire que A est diagonalisable.

Exemple

Déterminer les éléments propres de la matrice $A = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & 0 & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$.

• Valeurs propres

Facilement, $\operatorname{rg}(A) = 2$ donc 0 est valeur propre de multiplicité supérieure ou égale à $n - 2$ a priori. Il reste donc à trouver les deux autres valeurs propres λ et μ dans $\mathbb{C}$ (éventuellement égales).

Tout d'abord, $\operatorname{tr}(A) = \lambda + \mu = 2$. On calcule ensuite :

$$A^2 = \begin{pmatrix} n & 2 & \cdots & 2 & n \\ 2 & 2 & \cdots & 2 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 2 & \cdots & 2 & 2 \\ n & 2 & \cdots & 2 & n \end{pmatrix}.$$

✓ *On peut pour cela utiliser la formule du produit matriciel, mais il est plus rapide d'utiliser l'endomorphisme canoniquement associé à A.*

Donc, en utilisant la remarque faite auparavant :

$$\lambda^2 + \mu^2 = \operatorname{tr}(A^2) = 2(n-2) + 2n = 4n - 4\,.$$

On doit donc résoudre le système $\begin{cases} \lambda + \mu & = 2 \\ \lambda^2 + \mu^2 & = 4n - 4\,. \end{cases}$

Puisque $2\lambda\mu = (\lambda+\mu)^2 - (\lambda^2+\mu^2)$ *on a aussi* $\lambda\mu = 4 - 2n$ *et donc* λ *et* μ *sont racines de l'équation* $x^2 - 2x + 4 - 2n = 0$ *(en utilisant les relations entre somme et produit des racines et les coefficients). Ainsi :*

$$\{\lambda,\mu\} = \left\{1 + \sqrt{2n-3}, 1 - \sqrt{2n-3}\right\}.$$

Les valeurs propres de A sont donc finalement :

$$0 \text{ (d'ordre } n-2\text{)},\ 1 + \sqrt{2n-3} \ \text{ et } \ 1 - \sqrt{2n-3} \text{ (simples).}$$

• Sous-espaces propres

– *Le sous-espace propre associé à la valeur propre 0 ne pose pas de problème : en notant* $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$ *on a :*

$$AX = 0 \iff \begin{cases} x_1 + \cdots + x_n & = 0 \\ x_1 + x_n & = 0 \end{cases}.$$

Il s'agit de l'intersection de deux hyperplans distincts, donc c'est un sous-espace vectoriel de dimension $n-2$ (ce que l'on savait déjà), et il n'est pas très difficile d'en trouver une base si besoin est.

– *Pour trouver les autres sous-espaces propres, la résolution du système $AX = \lambda X$ avec $\lambda = 1 \pm \sqrt{2n-3}$ n'est pas très enthousiasmante... Nous allons donc utiliser une propriété importante et trop méconnue, citée et démontrée en* **[S4.1]** *et rappelée ci-dessous.*

> Si x est un vecteur propre d'un endomorphisme u associé à une valeur propre non nulle, alors $x \in \operatorname{Im} u$.

En notant $(e_1,\dots,e_n)$ la base canonique de $\mathbb{R}^n$ et a l'endomorphisme canoniquement associé à A, on voit que l'image de a est le sous-espace vectoriel engendré par les vecteurs $\sum_{i=1}^{n} e_i$ *et* $e_1 + e_n$, *donc aussi par* $u = \sum_{i=2}^{n-1} e_i$ *et* $v = e_1 + e_n$.

✓ *N'oubliez pas que l'image d'une application linéaire est le sous-espace vectoriel engendré par les images des vecteurs d'une base, ici représentées par les colonnes de la matrice.*

On cherche donc les autres vecteurs propres sous la forme $x = \alpha u + \beta v$. *Or* $a(u) = \sum_{i=2}^{n-1} a(e_i) = (n-2)v$ *et* $a(v) = a(e_1) + a(e_n) = 2\sum_{i=1}^{n} e_i = 2u + 2v$, *donc on a :*

$$\begin{aligned} a(x) = \lambda x &\iff \alpha a(u) + \beta a(v) = \lambda(\alpha u + \beta v) \\ &\iff \alpha(n-2)v + \beta(2u+2v) = \lambda(\alpha u + \beta v) \\ &\iff \begin{cases} 2\beta = \lambda\alpha \\ (n-2)\alpha + 2\beta = \lambda\beta \end{cases} \iff \begin{cases} \lambda\alpha - 2\beta = 0 \\ (n-2)\alpha + (2-\lambda)\beta = 0 \end{cases} \end{aligned}$$

Le déterminant de ce système est $\begin{vmatrix} \lambda & -2 \\ (n-2) & (2-\lambda) \end{vmatrix} = -\lambda^2 + 2\lambda + 2n - 4$. *Ce déterminant est nul puisque* λ *est une valeur propre (c'est rassurant, sinon la seule solution eût été* $\alpha = \beta = 0$ *!). Les deux équations sont donc équivalentes, et on en déduit qu'une base de la droite propre associée à* λ *est le vecteur de coordonnées* $(2,\lambda)$ *dans la base* (u,v) *de* $\operatorname{Im} a$, *c'est-à-dire le vecteur* $2u + \lambda v$, *de coordonnées* $(\lambda,2,2,\dots,2,\lambda)$ *dans la base canonique. Cela vaut pour les deux valeurs de* λ , $1 \pm \sqrt{2n-3}$.

SF4.6 Repérer des vecteurs propres évidents

Dans beaucoup de problèmes de recherches de vecteurs propres d'une matrice, certaines combinaisons linéaires sur les colonnes de cette matrice font apparaître des vecteurs propres « évidents ». Cela peut simplifier la suite des calculs (comme dans les exemples pages 293, 295 et 298), voire complètement résoudre l'exercice, comme dans l'exemple suivant.

Exemple

Soient a, $b \in \mathbb{C}$, $b \neq 0$. Déterminer les éléments propres de la matrice carrée d'ordre $2n$ donnée par : $A_n(a,b) = \begin{pmatrix} a & & & & b \\ & \ddots & & \iddots & \\ 0 & & & & 0 \\ & \iddots & & \ddots & \\ b & & & & a \end{pmatrix}$.

En notant $e_1,\dots,e_{2n}$ *les vecteurs de la base canonique, il est facile de vérifier que les vecteurs :*

$$e_1 + e_{2n} \ , \ e_2 + e_{2n-1} \ , \ e_n + e_{n+1}$$

sont n *vecteurs propres linéairement indépendants associés à la valeur propre* $a+b$, *et que les vecteurs :*

$$e_1 - e_{2n} \ , \ e_2 - e_{2n-1} \ , \ e_n - e_{n+1}$$

sont n *vecteurs propres linéairement indépendants associés à la valeur propre* $a-b$.
On a donc trouvé deux sous-espaces propres de dimension n, *on a ainsi trouvé tous les éléments propres et* A *est diagonalisable.*

✓ *On aurait pu ici facilement calculer le polynôme caractéristique (par récurrence ; il est égal à* $\left((X-a)^2 - b^2\right)^n$*). Mais cela eût été une perte de temps.*

SF4.7 Utiliser des sous-espaces stables

L'idée de cette méthode est la suivante : si l'espace vectoriel est somme directe de sous-espaces F_i stables par un endomorphisme u, les éléments propres de u s'obtiendront en cherchant les éléments propres des endomorphismes induits par u sur les F_i. On se ramène donc au même problème en dimension inférieure.

Cette méthode a déjà été utilisée dans un exemple de **[SF4.5]** page 295, en considérant l'image de l'endomorphisme (qui est évidemment stable). Voici un autre exemple.

Exemple

Soient a, b, c trois nombres complexes. La matrice $A = \begin{pmatrix} a & b & b & b & c \\ b & a & b & b & b \\ b & b & a & b & b \\ b & b & b & a & b \\ c & b & b & b & a \end{pmatrix}$ est-elle diagonalisable ?

Notons $(e_1, \dots, e_5)$ la base canonique de $\mathbb{C}^5$ et u l'endomorphisme de $\mathbb{C}^5$ canoniquement associé à A.

Déjà on remarque que :

$$u(e_1 - e_5) = (a - c)(e_1 - e_5)\ ,$$
$$u(e_2 - e_4) = (a - b)(e_2 - e_4)\ ,$$
$$u(e_3 - e_4) = (a - b)(e_3 - e_4)\,.$$

Nous avons ainsi déterminé trois vecteurs propres linéairement indépendants, c'est-à-dire trois droites stables. Pour trouver un plan stable, on peut s'inspirer de ce que nous savons dans le cas réel : en effet, lorsque a, b, c sont réels, la matrice A est symétrique réelle donc on sait que ses sous-espaces propres sont orthogonaux ; or dans $\mathbb{R}^5$ les vecteurs $e_2 + e_3 + e_4$ et $e_1 + e_5$ seraient orthogonaux aux trois vecteurs précédents. Il est donc légitime de calculer :

$$u(e_1 + e_5) = (a + c)(e_1 + e_5) + 2b(e_2 + e_3 + e_5)\,,$$
$$u(e_2 + e_3 + e_4) = 3b(e_1 + e_5) + (a + 2b)(e_2 + e_3 + e_4)\,.$$

Cela montre que le plan de base $(e_1 + e_5, e_2 + e_3 + e_4)$ est stable par u, et que la matrice dans cette base de l'endomorphisme induit par u est :

$$A' = \begin{pmatrix} a + c & 3b \\ 2b & a + 2b \end{pmatrix} = a\mathrm{I}_2 + \begin{pmatrix} c & 3b \\ 2b & 2b \end{pmatrix}.$$

La matrice A est diagonalisable si et seulement si A' l'est, ce qui équivaut aussi à la diagonalisabilité de $A'' = \begin{pmatrix} c & 3b \\ 2b & 2b \end{pmatrix}$. Le polynôme caractéristique de cette matrice est $\chi_{A''} = X^2 - (c + 2b)X + 2bc - 6b^2$, de discriminant $(c - 2b)^2 + 24b^2$. Il y a donc deux cas.

- *Si $(c - 2b)^2 + 24b^2 \neq 0$, A'' possède deux valeurs propres distinctes donc est diagonalisable.*
- *Si $(c - 2b)^2 + 24b^2 = 0$ (on est dans $\mathbb{C}$), A'' possède une valeur propre double, donc n'est diagonalisable que si elle est scalaire, ce qui correspond ici à $b = c = 0$.*

En conclusion : A est diagonalisable si et seulement si $(c-2b)^2+24b^2\neq 0$ ou $b=c=0$.

La notion de sous-espace stable par un endomorphisme est importante, et sert dans de nombreux exercices plus théoriques, en utilisant les résultats exposés en **[S4.11]**.

Exemple

Soient u et v deux endomorphismes diagonalisables d'un espace vectoriel E de dimension finie.
On suppose que u et v commutent. Montrer qu'il existe une base de E formée de vecteurs propres communs à u et à v.

Soit λ une valeur propre de u. Puisque u et v commutent, il en est de même de $u-\lambda\mathrm{Id}_E$ et de v, donc v laisse stable chaque sous-espace propre $E_\lambda(u)=\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$ (cf. **[S4.6]***).*
v étant diagonalisable, il en est de même de l'endomorphisme induit par v sur $E_\lambda(u)$, donc il existe une base $\mathscr{B}_\lambda$ de $E_\lambda(u)$ formée de vecteurs propres de v, qui sont aussi des vecteurs propres de u pour la valeur propre λ.
Puisque u est diagonalisable, les sous-espaces propres $E_\lambda(u)$ pour $\lambda\in\mathrm{Sp}\,u$ sont supplémentaires, donc $\bigcup_{\lambda\in\mathrm{Sp}\,u}\mathscr{B}_\lambda$ est une base de E, qui répond à la question.

SF4.8 Utiliser des matrices plus simples

S i une matrice A peut s'exprimer à l'aide d'une matrice B plus simple, la réduction de B peut aider à celle de A. Citons deux cas typiques.

- Lorsque A est de la forme $A=aI_n+B$.
  Dans ce cas, si X est un vecteur colonne et λ un scalaire, on a :
  $$AX=\lambda X \iff BX=(\lambda-a)X$$
  donc les valeurs propres de A seront de la forme $a+\mu$ avec μ valeur propre de B, et les sous-espaces propres seront les mêmes.
- Lorsque A est de la forme $A=P(B)$ avec P polynôme.
  Dans ce cas, si l'on a trouvé une matrice $Q\in\mathrm{GL}_n(\mathbb{K})$ et une matrice B' diagonale ou triangulaire supérieure telles que $B=QB'Q^{-1}$, puisque $B^k=QB'^kQ^{-1}$ pour tout entier k, on a :
  $$A=P(B)=QP(B')Q^{-1}\,,$$
  c'est-à-dire que A est semblable à la matrice $P(B')$ qui est elle aussi diagonale ou triangulaire supérieure. En particulier, les valeurs propres de A sont les éléments de la diagonale de $P(B')$.

Exemple

En utilisant la matrice $J_n\in\mathcal{M}_n(\mathbb{R})$ dont tous les éléments sont égaux à 1, diagonaliser la matrice $A=\begin{pmatrix}a & b & \dots & b\\ b & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & b\\ b & \dots & b & a\end{pmatrix}$.

- *On sait que* $J_n^2 = nJ_n$ *(le calcul a été fait en* **[SF3.6]** *page 253).*
  On en déduit que le polynôme $X^2 - nX$ *est annulateur de* J_n *donc que* $\mathrm{Sp}\, J_n \subset \{0,n\}$.
  - *Si* X *est le vecteur colonne* ${}^{\mathrm{t}}(x_1 \quad \dots \quad x_n)$, *le système* $J_nX = 0$ *équivaut à la seule équation* $x_1 + \cdots + x_n = 0$ *: le sous-espace propre* $E_0(J_n)$ *associé à la valeur propre* 0 *est donc l'hyperplan d'équation* $x_1 + \cdots + x_n = 0$.
  - *De plus, si* $V = {}^{\mathrm{t}}(1 \quad 1 \quad \dots \quad 1)$ *on a facilement* $J_nV = nV$ *donc* V *est un vecteur propre associé à la valeur propre* n. *La droite vectorielle engendrée par* V *est donc a priori incluse dans le sous-espace propre* $E_1(J_n)$. *Mais ce sous-espace propre étant en somme directe avec* $E_0(J_n)$ *(cf.* **[S4.2]***) ne peut être de dimension supérieure à* 1 *; c'est donc exactement la droite* $\mathrm{Vect}(V)$.

  Puisque la somme des dimensions des sous-espaces propres est égale à n, *la matrice* J *est diagonalisable.*

  ✓ *On pouvait aussi remarquer que* J_n *est symétrique réelle, donc diagonalisable, et ses sous-espaces propres sont orthogonaux. Ainsi, une fois trouvé le 1er sous-espace propre, on pouvait conclure directement.*

- *On remarque alors que l'on a :* $A = (a-b)\mathrm{I}_n + bJ_n$. *Donc :*
$$AX = \lambda X \iff bJ_nX = (\lambda - a + b)X\,.$$
  Si $b = 0$, *la matrice* A *est déjà diagonale ; le problème est résolu.*
  Sinon, le système ci-dessus équivaut à $J_nX = \dfrac{\lambda - a + b}{b}X$. *On déduit alors directement de l'étude de* J_n *que :*
  - $\lambda = a - b$ *est valeur propre de* A *; le sous-espace propre associé est l'hyperplan d'équation* $x_1 + \cdots + x_n = 0$ *;*
  - $\dfrac{\lambda - a + b}{b} = n \iff \lambda = a + (n-1)b$ *donc* $\lambda = a + (n-1)b$ *est valeur propre de* A, *le sous-espace propre associé étant la droite vectorielle de base* ${}^{\mathrm{t}}(1 \quad 1 \quad \dots \quad 1)$.

  On en déduit, pour des raisons de dimension, que A *est diagonalisable.*

Exemple

1) Soit n un entier $\geqslant 2$. Déterminer les éléments propres de la matrice :
$$J_n = \begin{pmatrix} 0 & 1 & 0 & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & 1 & 0 \\ & 0 & & \ddots & 1 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})\,.$$

2) Calculer les puissances de J_n et en déduire les éléments propres de la matrice :
$$C(a_0,\dots,a_{n-1}) = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})\,.$$

1) *Soit* $X = {}^{\mathrm{t}}\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$, *non nul, et* $\lambda \in \mathbb{C}$ *tels que que* $J_n X = \lambda X$. *On a alors le système :*

$$x_2 = \lambda x_1, \quad x_3 = \lambda x_2, \quad \dots \quad x_n = \lambda x_{n-1}, \quad x_1 = \lambda x_n\,.$$

On en déduit $x_1 = \lambda^n x_1$ *d'où* $\lambda^n = 1$ *puisque* $x_1 \neq 0$ *(en effet,* $x_1 = 0$ *impliquerait* $X = 0$, *ce qui est exclu pour un vecteur propre). Les valeurs propres de* J_n *sont donc incluses dans l'ensemble des racines n-ièmes de l'unité.*

> ✓ *On aurait d'ailleurs pu remarquer que* $J_n^n = \mathrm{I}_n$ *c'est-à-dire que* $X^n - 1$ *est annulateur de* J_n.

Réciproquement, si ω *est une racine n-ème de l'unité quelconque, le système précédent montre que le vecteur* ${}^{\mathrm{t}}\begin{pmatrix} 1 & \omega & \omega^2 & \dots & \omega^{n-1} \end{pmatrix}$ *est un vecteur propre de* J_n.

La matrice J_n, *possédant n valeurs propres distinctes est diagonalisable. Posons* $\omega = \mathrm{e}^{\frac{2\mathrm{i}\pi}{n}}$, *de sorte que les racines n-ièmes de l'unité sont les complexes* ω^k *pour* $0 \leqslant k \leqslant n-1$.

On a donc $J_n = PDP^{-1}$ *où* $D = \mathrm{diag}(1,\omega,\dots,\omega^{n-1})$ *et où* P *est la matrice dont la* j^e *colonne est* ${}^{\mathrm{t}}\begin{pmatrix} 1 & \omega^{j-1} & \dots & \omega^{(j-1)(n-1)} \end{pmatrix}$.

2) • *Pour calculer les puissances de* J, *la méthode la plus élégante consiste à utiliser les applications linéaires. Si* E *est un espace vectoriel de dimension* n *rapporté à une base* $\mathscr{B} = (e_1,\dots,e_n)$, J *est la matrice dans cette base de l'endomorphisme* u *tel que :*

$$u(e_1) = e_n\,,\ u(e_2) = e_1\,,\ u(e_3) = e_2\ \dots\ u(e_n) = e_{n-1}\,,$$

donc qui transforme la base $\mathscr{B}$ *en la base* $\sigma(\mathscr{B}) = (e_n,e_1,e_2,\dots,e_{n-1})$ *obtenue après une permutation circulaire* σ. *Pour tout entier* $k \in [\![1\,;n-1]\!]$ J_n^k *sera donc la matrice de l'endomorphisme* u^k *qui transforme la base* $\mathscr{B}$ *en la base* $\sigma^k(\mathscr{B}) = (e_{n-k+1},\dots,e_n,e_1,\dots,e_{n-k})$ *où le vecteur* e_n *est à la* k^e *place.*

Avec la notation de l'énoncé, il s'agit de la matrice $C(0,0,\dots,0,1,0,\dots,0)$, *le* 1 *étant à la* $(k+1)^e$ *place.*

> ✓ *On constatera, une fois encore, que l'interprétation d'une matrice à l'aide des applications linéaires permet un calcul bien plus facile que l'utilisation des seules formules du produit matriciel.*

• *On a donc :* $\forall\,(a_0,\dots,a_{n-1}) \in \mathbb{C}^n$, $C(a_0,\dots,a_{n-1}) = \displaystyle\sum_{k=0}^{n-1} a_k J_n^k$.

Cela montre au passage que l'ensemble des matrices de la forme $C(a_0,\dots,a_{n-1})$ *est le sous-espace vectoriel de* $\mathcal{M}_n(\mathbb{K})$*engendré par les matrices* $\mathrm{I}_n, J_n, \dots, J_n^{n-1}$. *De plus, avec les notations de la première question on a :*

$$C(a_0,\dots,a_{n-1}) = \sum_{k=0}^{n-1} a_k (PDP^{-1})^k = P\left(\sum_{k=0}^{n-1} a_k D^k\right) P^{-1}\,.$$

La matrice $\displaystyle\sum_{k=0}^{n-1} a_k D^k$ *étant diagonale, cela montre que* $C(a_0,\dots,a_{n-1})$ *est diagonalisable dans la même base que* J_n, *et que ses valeurs propres sont les* $\displaystyle\sum_{k=0}^{n-1} a_k \omega^k$ *lorsque* ω *décrit l'ensemble des racines* n^{es} *de l'unité.*

SF4.9 Utiliser le théorème spectral

Le théorème spectral permet de dire qu'une matrice symétrique réelle est diagonalisable. Mais on sait davantage.

- Les sous-espaces propres sont orthogonaux : cette propriété peut permettre, si l'on connait déjà certains sous-espaces propres, de trouver les autres.
- On peut trouver une base orthonormale de vecteurs propres, c'est-à-dire une matrice de passage P qui est orthogonale ; l'avantage d'une telle matrice est que l'on a $P^{-1} = {}^tP$, donc le calcul de l'inverse est immédiat.

Exemple

Diagonaliser la matrice $A = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}$.

A étant symétrique réelle est diagonalisable.

On remarque assez facilement que la matrice $A - 3\mathrm{I}_3$ est de rang 1 donc le noyau de $A - 3\mathrm{I}_3$ est de dimension 2. Cela signifie que 3 est valeur propre de A, d'ordre de multiplicité 2 (car la matrice étant diagonalisable, la dimension de chaque sous-espace propre est égale à l'ordre de multiplicité de la valeur propre correspondante).

Soit $V = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. $AV = 3V \iff x - y + z = 0$: le sous-espace propre associé à la valeur propre 3 est le plan d'équation $x - y + z = 0$.

Puisque la trace de A est égale à 3, la troisième valeur propre est égale à -3 ; le sous-espace propre associé est nécessairement la droite orthogonale au plan précédent, c'est-à-dire la droite de base (u) avec $u = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ (on l'a rendu unitaire).

Enfin, puisque le vecteur unitaire $v = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ appartient au plan trouvé précédemment, une base orthonormale de vecteurs propres sera formée des vecteurs u, v et $w = u \wedge v$, soit $w = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$.

En conclusion, on a $A = PDP^{-1}$ avec $D = diag(-3,3,3)$ et

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \quad (\text{avec } P^{-1} = {}^tP).$$

■ Comment étudier la diagonalisabilité d'une matrice ou d'un endomorphisme ?

SF4.10 Déterminer les éléments propres

Lorsque l'on a déterminé les éléments propres d'une matrice ou d'un endomorphisme, par l'une des méthodes vues auparavant, on peut généralement conclure quant à la diagonalisabilité. Il suffit pour cela d'étudier la dimension des sous-espaces propres, en utilisant l'un des théorèmes exposés en **[S4.5]**.

Nous avons déjà eu l'occasion d'appliquer cette méthode lors d'exemples précédents. Donnons néanmoins un autre exemple montrant comment on peut s'en servir dans un contexte différent.

Exemple

Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ une matrice diagonalisable, et $A \in \mathcal{M}_n(\mathbb{K})$. On considère la matrice $B \in \mathcal{M}_{2n}(\mathbb{K})$ définie par blocs par : $B = \begin{bmatrix} aA & bA \\ cA & dA \end{bmatrix}$.

Montrer que : A diagonalisable $\implies$ B diagonalisable.

Nous allons déterminer les sous-espaces propres de B à l'aide de ceux de A, en faisant des calculs par blocs.

- *Si 0 est valeur propre de A, soit alors $X \in \mathcal{M}_{n,1}(\mathbb{K})$, non nul, tel que $AX = 0$. Alors pour tous réels α, β, $B\begin{bmatrix} \alpha X \\ \beta X \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, donc si $(\alpha,\beta) \neq (0,0)$ $\begin{bmatrix} \alpha X \\ \beta X \end{bmatrix}$ est un vecteur propre de B associés à la valeur propre 0.*
- *Soit maintenant λ une valeur propre non nulle de A, et $X \in \mathcal{M}_{n,1}(\mathbb{K})$ un vecteur propre associé. Par analogie avec le cas précédent, nous allons chercher des vecteurs propres de B sous la forme $Y = \begin{bmatrix} \alpha X \\ \beta X \end{bmatrix}$.*

$$BY = \mu Y \iff \begin{bmatrix} aA & bA \\ cA & dA \end{bmatrix} \begin{bmatrix} \alpha X \\ \beta X \end{bmatrix} = \mu \begin{bmatrix} \alpha X \\ \beta X \end{bmatrix}$$

$$\iff \begin{cases} \lambda(a\alpha + b\beta) &= \mu\alpha \\ \lambda(c\alpha + b\beta) &= \mu\beta \end{cases} \quad \text{puisque } X \neq 0 \text{ vérifie } AX = \lambda X$$

$$\iff \begin{cases} a\alpha + b\beta &= \dfrac{\mu}{\lambda}\alpha \\ c\alpha + b\beta &= \dfrac{\mu}{\lambda}\beta \end{cases} \quad \text{puisque } \lambda \neq 0$$

$$\iff \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{ est vecteur propre de } M \text{ pour la valeur propre } \frac{\mu}{\lambda}.$$

Or par hypothèse on dispose d'une base $\left(\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}\right)$ de vecteurs propres de M, pour des valeurs propres que nous noterons respectivement ν_1, ν_2. D'après le calcul précédent on a :

$$B\begin{bmatrix} \alpha_1 X \\ \beta_1 X \end{bmatrix} = \lambda\nu_1 \begin{bmatrix} \alpha_1 X \\ \beta_1 X \end{bmatrix} \quad \text{et} \quad B\begin{bmatrix} \alpha_2 X \\ \beta_2 X \end{bmatrix} = \lambda\nu_2 \begin{bmatrix} \alpha_2 X \\ \beta_2 X \end{bmatrix}$$

et les vecteurs $\begin{bmatrix}\alpha_1 X\\ \beta_1 X\end{bmatrix}$ *et* $\begin{bmatrix}\alpha_2 X\\ \beta_2 X\end{bmatrix}$ *sont donc deux vecteurs propres de* B.

- *A étant diagonalisable, il existe une base $(X_1,\dots,X_n)$ de vecteurs propres de A. D'après les deux cas étudiés précédemment, la famille formée des $2n$ vecteurs* $\begin{bmatrix}\alpha_1 X_1\\ \beta_1 X_1\end{bmatrix},\dots,\begin{bmatrix}\alpha_1 X_n\\ \beta_1 X_n\end{bmatrix},\begin{bmatrix}\alpha_2 X_1\\ \beta_2 X_1\end{bmatrix},\dots,\begin{bmatrix}\alpha_2 X_n\\ \beta_2 X_n\end{bmatrix}$ *est une famille de vecteurs propres de B.*

  Montrons que cette famille est libre : s'il existe des scalaires λ_i et μ_i tels que :

$$\sum_{i=1}^{n}\lambda_i\begin{bmatrix}\alpha_1 X_i\\ \beta_1 X_i\end{bmatrix}+\mu_i\begin{bmatrix}\alpha_2 X_i\\ \beta_2 X_i\end{bmatrix}=\begin{bmatrix}0_n\\ 0_n\end{bmatrix},$$

  on a :

$$\sum_{i=1}^{n}(\lambda_i\alpha_1+\mu_i\alpha_2)X_i=\sum_{i=1}^{n}(\lambda_i\beta_1+\mu_i\beta_2)X_i=0_n,$$

  et, la famille $(X_i)_{1\leqslant i\leqslant n}$ étant libre, on en déduit :

$$\lambda_i\alpha_1+\mu_i\alpha_2=\lambda_i\beta_1+\mu_i\beta_2=0 \text{ pour tout } i$$

  soit encore

$$\lambda_i\begin{pmatrix}\alpha_1\\ \beta_1\end{pmatrix}+\mu_i\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix} \text{ pour tout } i,$$

  et la famille $\left(\begin{pmatrix}\alpha_1\\ \beta_1\end{pmatrix},\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix}\right)$ *étant libre, il vient $\lambda_i=\mu_i=0$ pour tout i.*

- *On a donc trouvé une base de $\mathcal{M}_{2n,1}(\mathbb{K})$ formée de vecteurs propres de B, ce qui prouve que B est diagonalisable.*

  Au passage, on a aussi démontré que les valeurs propres de B sont les produits des valeurs propres de M et de celles de A.

SF4.11 Utiliser le théorème du rang

Pour déterminer la dimension des sous-espaces propres, on peut utiliser le théorème du rang. Plus précisément, si u est un endomorphisme d'un espace vectoriel E de dimension finie, on a :

$$\dim E_\lambda(u)=\dim E-\operatorname{rg}(u-\lambda\mathrm{Id}_E).$$

Dans certains cas, le calcul de $\operatorname{rg}(u-\lambda\mathrm{Id}_E)$ peut être plus rapide que la résolution de systèmes compliqués pour déterminer exactement $E_\lambda(u)$.

Exemple

Donner une condition nécessaire et suffisante portant sur les réels a, b, c, d, e et f pour que la matrice $A=\begin{pmatrix}1&a&b&c\\0&1&d&e\\0&0&2&f\\0&0&0&2\end{pmatrix}$ soit diagonalisable.

La matrice A étant triangulaire supérieure, ses valeurs propres sont ses éléments diagonaux. Ses valeurs propres sont donc 1 et 2, chacune étant double.

Pour que A soit diagonalisable, il faut et il suffit que :

$$\dim E_1(A) = 2 \quad \text{et} \quad \dim E_2(A) = 2\,,$$

ce qui équivaut, grâce au théorème du rang, à :

$$\operatorname{rg}(A - \mathrm{I}_4) = \operatorname{rg}(A - 2\mathrm{I}_4) = 2\,.$$

Or $A - \mathrm{I}_4 = \begin{pmatrix} 0 & a & b & c \\ 0 & 0 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{pmatrix}$, *donc :* $\operatorname{rg}(A - \mathrm{I}_4) = 2 \iff a = 0\,.$

Et $A - 2\mathrm{I}_4 = \begin{pmatrix} -1 & a & b & c \\ 0 & -1 & d & e \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & 0 \end{pmatrix}$, *donc :* $\operatorname{rg}(A - 2\mathrm{I}_4) = 2 \iff f = 0\,.$

En conclusion : A est diagonalisable si et seulement si $a = f = 0$.

SF4.12 Utiliser un polynôme annulateur

Il s'agit ici, sans forcément chercher les sous-espaces propres, d'utiliser les résultats rappelés en **[S4.10]**, et en particulier celui-ci.

> Un endomorphisme u d'un $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si il possède un polynôme annulateur scindé dans $\mathbb{K}[X]$ et à racines simples.

Il s'agit certainement de la méthode la plus efficace dans le cas d'exercices un peu théoriques. La difficulté est qu'elle n'apparaît pas toujours directement ; c'est pourquoi nous donnons plusieurs exemples de son utilisation (voir aussi un exemple de **[SF4.2]** page 290).

Exemple

Démontrer qu'une matrice carrée de rang 1 est diagonalisable si et seulement si sa trace n'est pas nulle.

Soit $A \in \mathcal{M}_n(\mathbb{R})$, de rang 1 (il faut supposer $n \geqslant 2$ bien sûr). Alors toutes les colonnes de A sont proportionnelles à une même colonne $C = {}^{\mathrm{t}}(c_1 \quad \dots \quad c_n)$. Par blocs, A s'écrit donc sous la forme :

$$A = \begin{bmatrix} \ell_1 C & \ell_2 C & \cdots & \ell_n C \end{bmatrix}.$$

En considérant la matrice ligne $L = \begin{pmatrix} \ell_1 & \ell_2 & \cdots & \ell_n \end{pmatrix}$, on a donc $A = CL$.

On a alors $A^2 = CLCL$; or $LC = \sum_{i=1}^{n} \ell_i c_i = \operatorname{tr}(A)$ est un scalaire, donc :

$$A^2 = \left(\sum_{i=1}^{n} \ell_i c_i \right) CL = \operatorname{tr}(A) A\,.$$

- *Si $\operatorname{tr}(A) \neq 0$, A annule le polynôme $X^2 - \operatorname{tr}(A)X$ qui est scindé à racines simples, donc A est diagonalisable.*
- *Si $\operatorname{tr}(A) = 0$ alors $A^2 = 0$ et 0 est donc la seule valeur propre de A ; si A était diagonalisable, elle serait semblable, donc égale, à la matrice nulle, ce qui n'est pas le cas.*

On a donc bien prouvé l'équivalence demandée.

Exemple

Soit G un sous-groupe fini du groupe $(\mathrm{GL}_n(\mathbb{R}),\times)$. Montrer que :

1) $\forall\, M \in G,\ \det M = \pm 1$ et $|\mathrm{tr}(M)| \leqslant n$.

2) $\forall\, M \in G,\ \mathrm{tr}(M) = \mathrm{tr}(M^{-1})$.

Soit m le cardinal de G. Alors pour toute matrice $A \in G$, on a $A^m = \mathrm{I}_n$, c'est-à-dire que le polynôme $X^m - 1$ est annulateur de A.
Ce polynôme étant scindé à racines simples dans $\mathbb{C}$, A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$, et ses valeurs propres complexes λ_i sont des racines m^{es} de l'unité.

1) $\det M$ *est le produit de ces valeurs propres, donc est de module* 1. *Puisqu'il s'agit d'un réel, on a* $\det M = \pm 1$.
$\mathrm{tr}(M)$ *est la somme de ces valeurs propres, donc d'après l'inégalité triangulaire,* $|\mathrm{tr}(M)| \leqslant \sum_{i=1}^{n} |\lambda_i| = n$.

2) *Puisque M est semblable (dans $\mathcal{M}_n(\mathbb{C})$) à* $\mathrm{diag}(\lambda_1,\dots,\lambda_n)$, M^{-1} *est semblable à :*

$$\mathrm{diag}\left(\frac{1}{\lambda_1},\dots,\frac{1}{\lambda_n}\right) = \mathrm{diag}(\overline{\lambda_1},\dots,\overline{\lambda_n}) \quad (\textit{car } |\lambda_i| = 1 \Longrightarrow \frac{1}{\lambda_i} = \overline{\lambda_i}).$$

Donc $\mathrm{tr}(M^{-1}) = \overline{\mathrm{tr}(M)} = \mathrm{tr}(M)$, *puisque c'est un nombre réel.*

SF4.13 Utiliser le théorème de décomposition des noyaux

Exemple

Soit E un $\mathbb{C}$-espace vectoriel de dimension n, et u un endomorphisme de E.

Montrer que u est diagonalisable si et seulement si u^2 est diagonalisable et $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$.

- *Supposons u diagonalisable. Il existe donc une base de E dans laquelle la matrice de u est* $\mathrm{diag}(\lambda_1,\dots,\lambda_r,0,\dots,0)$, *où $\lambda_1,\dots,\lambda_r$ sont les valeurs propres non nulles de u, et $r = \mathrm{rg}\, u$.*
Dans cette même base, la matrice de u^2 sera donc $\mathrm{diag}(\lambda_1^2,\dots,\lambda_r^2,0,\dots,0)$, *ce qui montre que u^2 est diagonalisable et que* $\mathrm{rg}\, u^2 = r$. *Puisque* $\mathrm{rg}\, u^2 = \mathrm{rg}\, u$, *on a d'après le théorème du rang* $\dim \mathrm{Ker}\, u^2 = \dim \mathrm{Ker}\, u$ *et puisque* $\mathrm{Ker}\, u \subset \mathrm{Ker}\, u^2$, *on en déduit l'égalité* $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$.
- *Réciproquement supposons u^2 diagonalisable et* $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$. *Notons $\mu_1,\dots,\mu_p$ ses valeurs propres distinctes non nulles. E étant somme directe des sous-espaces propres de u^2 on a :*

$$E = (\mathrm{Ker}\, u^2) \bigoplus \left(\bigoplus_{i=1}^{p} \mathrm{Ker}(u^2 - \mu_i \mathrm{Id}_E) \right)$$

(si u est bijectif, $\mathrm{Ker}\, u^2 = \{0\}$, *sinon c'est un sous-espace propre).*
Chaque complexe μ_i étant non nul possède deux racines complexes distinctes $\pm\lambda_i$; d'après le théorème de décomposition des noyaux, on a

$$\mathrm{Ker}(u^2 - \lambda_i^2 \mathrm{Id}_E) = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E) \oplus \mathrm{Ker}(u + \lambda_i \mathrm{Id}_E)$$

puisque les polynômes $X-\lambda_i$ et $X+\lambda_i$ sont premiers entre eux. On en déduit :

$$E = (\operatorname{Ker} u^2) \bigoplus \left(\bigoplus_{i=1}^{p} \operatorname{Ker}(u-\lambda_i \mathrm{Id}_E) \bigoplus \operatorname{Ker}(u+\lambda_i \mathrm{Id}_E) \right)$$
$$= (\operatorname{Ker} u) \bigoplus \left(\bigoplus_{i=1}^{p} \operatorname{Ker}(u-\lambda_i \mathrm{Id}_E) \bigoplus \operatorname{Ker}(u+\lambda_i \mathrm{Id}_E) \right)$$

donc E est somme directe de sous-espaces propres de u, ce qui prouve que u est diagonalisable.

■ Comment utiliser la trigonalisation ?

Il y a très peu de choses à savoir concernant la trigonalisation dans le programme officiel. Le seul résultat est le suivant.

> Un endomorphisme d'un $\mathbb{K}$-espace vectoriel est trigonalisable si et seulement si il possède un polynôme annulateur scindé dans $\mathbb{K}[X]$ (ce qui équivaut à dire que son polynôme caractéristique est scindé).
> En particulier, tout endomorphisme d'un $\mathbb{C}$-espace vectoriel est trigonalisable.

SF4.14 Utiliser la trigonalisabilité

Le théorème rappelé ci-dessus a des conséquences importantes. Nous allons en citer deux, très classiques.

Exemple

Soit E un $\mathbb{C}$-espace vectoriel de dimension finie $n \geqslant 1$.

Montrer qu'un endomorphisme $u \in \mathscr{L}(E)$ est nilpotent si et seulement si sa seule valeur propre est 0.

- *Si u est nilpotent, il existe $p \in \mathbb{N}^*$ tel que $u^p = 0_{\mathscr{L}(E)}$. Cela signifie que le polynôme X^p est annulateur de u. Puisque le spectre d'un endomorphisme est inclus dans l'ensemble des racines d'un polynôme annulateur, on en déduit :* $\operatorname{Sp} u \subset \{0\}$.
  De plus, un endomorphisme nilpotent ne peut être injectif (car, par exemple, u injectif implique u^p injectif pour tout entier p), donc 0 est nécessairement valeur propre de u et finalement :
  $$\operatorname{Sp} u = \{0\}.$$
- *Réciproquement, soit $u \in \mathscr{L}(E)$ tel que $\operatorname{Sp} u = \{0\}$. u étant trigonalisable, il existe une base $\mathscr{B} = (e_1, \ldots, e_n)$ telle que la matrice de u dans cette base soit triangulaire supérieure stricte (c'est-à-dire avec des zéros sur la diagonale). Cela signifie que, si l'on note :*
  $$E_0 = \{0\} \text{ et } E_k = \operatorname{Vect}(e_1, \ldots, e_k) \text{ pour tout } k \in [\![1\,;n]\!],$$
  alors on a $u(E_k) \subset E_{k-1}$ pour tout $k \in [\![1\,;n]\!]$.
  Il en résulte immédiatement que $u^n(E) = u^n(E_n) = E_0 = \{0\}$ donc $u^n = 0$, et u est nilpotent.

✓ On a (re)démontré au passage une propriété importante : si u est nilpotent, son indice de nilpotence, c'est-à-dire le plus petit entier p tel que $u^p = 0$, est nécessairement inférieur ou égal à n.

Exemple

Soit u un endomorphisme d'un $\mathbb{C}$-espace vectoriel de dimension finie n.
On suppose que : $\forall\, k \in [\![1\,;n]\!],\ \operatorname{tr} u^k = 0$.

Montrer que u est nilpotent

D'après l'exemple précédent, il suffit de montrer que $\operatorname{Sp} u = \{0\}$. *Raisonnons par l'absurde, et supposons que u possède des valeurs propres distinctes non nulles $\lambda_1, \ldots, \lambda_p$, d'ordre de multiplicité respectifs $m_1, \ldots, m_p$.*
On sait qu'il existe une base où la matrice de u est triangulaire supérieure, avec sur la diagonale les λ_i, chacun étant répété m_i fois, ainsi que des zéros (éventuels). Sur la diagonale de la matrice de u^k figurent donc les λ_i^k, chacun répété m_i fois donc on a :

$$\forall\, k \in \mathbb{N},\ \operatorname{tr} u^k = \sum_{i=1}^{p} m_i \lambda_i^k\,.$$

L'hypothèse de l'énoncé conduit donc au système de p équations :

$$\forall\, k \in [\![1\,;p]\!],\ \operatorname{tr} u^k = \sum_{i=1}^{p} m_i \lambda_i^k = 0\,.$$

qui peut s'écrire matriciellement : $\begin{pmatrix} \lambda_1 & \cdots & \lambda_p \\ \vdots & & \vdots \\ \lambda_1^p & \cdots & \lambda_p^p \end{pmatrix} \cdot \begin{pmatrix} m_1 \\ \vdots \\ m_p \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad (S)\,.$

Le déterminant de ce système vaut :

$$\begin{vmatrix} \lambda_1 & \cdots & \lambda_p \\ \vdots & & \vdots \\ \lambda_1^p & \cdots & \lambda_p^p \end{vmatrix} = \prod_{i=1}^{p} \lambda_i \begin{vmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ \lambda_i^{p-1} & \cdots & \lambda_p^{p-1} \end{vmatrix} = \left(\prod_{i=1}^{p} \lambda_i \right) V(\lambda_1, \ldots, \lambda_p)$$

où $V(\lambda_1, \ldots, \lambda_p)$ est un déterminant de Vandermonde (cf. **[S3.14]***). Puisque les λ_i sont non nuls et distincts, ce déterminant est non nul ; il en résulte que la seule solution du système (S) est $m_1 = \cdots = m_p = 0$, ce qui est absurde.*

En conclusion : la seule valeur propre de u est 0.

SF4.15 Trigonaliser une matrice

Pour trigonaliser effectivement une matrice A d'ordre n avec $n \leqslant 4$, on commence par chercher ses valeurs propres et ses sous-espaces propres ; en prenant déjà la réunion de bases de ces sous-espaces, on aura le début d'une future base de trigonalisation (cela donne la « partie diagonale » de la future matrice).

Exemple

Trigonaliser la matrice $A = \begin{pmatrix} 2 & 2 & -3 \\ 5 & 1 & -5 \\ -3 & 4 & 0 \end{pmatrix}$.

Les détails des calculs ne sont pas tous reproduits ici, mais nous vous invitons vivement à les faire par vous-même. On trouve, soit par le calcul du polynôme caractéristique, soit par une autre des méthodes déjà exposées, que 1 *est la seule valeur propre de* A.
On peut déjà dire que A *n'est pas diagonalisable car, si elle l'était, elle serait semblable à* I_3 *donc égale à* I_3.
On trouve facilement que le sous-espace propre associé à la valeur propre 1 *est la droite de base* $e_1 = (1,1,1)$.
Puisque χ_A *est scindé dans* $\mathbb{R}[X]$, *on sait qu'il existe une base* (e_1,e_2,e_3) *telle que la matrice dans cette base de l'endomorphisme* u *canoniquement associé à* A *soit égale à :* $T = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$ *avec* a, b, c *réels.*
Il reste donc à déterminer e_2 *et* e_3, *ainsi que les constantes* a, b, c, *les plus simples possibles.*

- *On doit déjà avoir* $u(e_2) = ae_1 + e_2$, *avec* a *non nul (sinon,* e_2 *serait colinéaire à* e_1). *En posant* $e_2 = (x,y,z)$ *cela équivaut au système :*
$$\begin{cases} 2x + 2y - 3z = a + x \\ 5x + y - 5z = a + y \\ -3x + 4y = a + z \end{cases} \text{ soit } \begin{cases} x + 2y - 3z = a \\ 5x - 5z = a \\ -3x + 4y - z = a \end{cases}.$$
On voit alors que l'on peut choisir $a = 5$ *(quitte à multiplier* x, y *et* z *par une même constante) ; ensuite on peut prendre (on a le choix)* $x = 0$, $y = 1$ *et* $z = -1$, *soit* $e_2 = (0,1,-1)$.
- *On cherche ensuite* e_3 *tel que* (e_1,e_2,e_3) *soit libre et* $u(e_3) = be_1 + ce_2 + e_3$ *ce qui donne le système :*
$$\begin{cases} 2x + 2y - 3z = b + x \\ 5x + y - 5z = b + c + y \\ -3x + 4y = b - c + z \end{cases} \text{ soit } \begin{cases} x + 2y - 3z = b \\ 5x - 5z = b + c \\ -3x + 4y - z = b - c \end{cases}.$$
On ne peut évidemment pas choisir $c = 0$ *puisqu'on obtiendrait le même système qu'avant ; si on choisit* $b = 0$ *et* $c = 10$ *on pourra prendre* $x = 2$, $y = -1$ *et* $z = 0$, *soit* $e_3 = (2,-1,0)$.

Pour conclure, on a $A = PTP^{-1}$ *avec :*
$$T = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 10 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

■ Comment calculer les puissances d'une matrice ?

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On cherche à calculer A^p pour $p \in \mathbb{N}$.

SF4.16 Utiliser un polynôme annulateur

Soit $P \in \mathbb{K}[X]$ un polynôme annulateur de A ; ce peut être son polynôme caractéristique (d'après le théorème de Cayley-Hamilton), ou son polynôme minimal (qui est, dans le cas où A est diagonalisable, égal à $\prod_{\lambda \in \mathrm{Sp}\, A} (X - \lambda)$).

Si d est le degré de P, la division euclidienne de X^p par P dans $\mathbb{K}[X]$ s'écrit : $X^p = PQ + R_p$, où le reste R_p est un polynôme de degré $\leqslant d-1$ à déterminer.

On a alors : $A^p = P(A)Q(A) + R_p(A) = R_p(A)$, ce qui permet d'exprimer A^p pour tout entier p en fonction de $\mathrm{I}_n, A, \ldots, A^{d-1}$.

✓ Il s'agit, à notre avis, de la méthode la plus rapide dans le cas général.

Exemple

Soit $A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$. Calculer A^n pour $n \in \mathbb{N}$.

On calcule le polynôme caractéristique $\chi_A(X) = \det(X\mathrm{I}_3 - A)$:

$$\chi_A(X) = \begin{vmatrix} X-1 & -3 & 0 \\ -3 & X+2 & 1 \\ 0 & 1 & X-1 \end{vmatrix} \underset{L_1 \leftarrow L_1 + 3L_3}{=} \begin{vmatrix} X-1 & 0 & 3(X-1) \\ -3 & X+2 & 1 \\ 0 & 1 & X-1 \end{vmatrix}$$

$$= (X-1)\begin{vmatrix} 1 & 0 & 3 \\ -3 & X+2 & 1 \\ 0 & 1 & X-1 \end{vmatrix} \underset{C_3 \leftarrow C_3 - 3C_1}{=} (X-1)\begin{vmatrix} 1 & 0 & 0 \\ -3 & X+2 & 10 \\ 0 & 1 & X-1 \end{vmatrix}$$

$$= (X-1)\begin{vmatrix} X+2 & 10 \\ 1 & X-1 \end{vmatrix} = (X-1)[(X+2)(X-1) - 10]$$

$$= (X-1)(X^2 + X - 12) = (X-1)(X-3)(X+4).$$

Ainsi : $\chi_A(X) = (X-1)(X-3)(X+4)$.

✓ *Il est en général maladroit d'utiliser la règle de Sarrus pour calculer le polynôme caractéristique d'une matrice carrée d'ordre 3, car cela donne le polynôme sous forme développée et non factorisée.*

D'après le cours sur les polynômes, la division euclidienne de X^n par χ_A s'écrit :

$$X^n = Q\chi_A + R_n \quad (1)$$

avec $\deg R_n < \deg \chi_A$, soit $\deg R_n \leqslant 2$. On pourra donc poser $R_n = a_n X^2 + b_n X + c_n$.
En appliquant la relation (1) en -4, 1 et 3 (ce sont les racines de χ_A), on obtient le système :

$$\left\{ \begin{array}{rcl} 16a_n - 4b_n + c_n &=& (-4)^n \\ a_n + b_n + c_n &=& 1 \\ 9a_n + 3b_n + c_n &=& 3^n \end{array} \right. \qquad (S_1)$$

Ce système est assez ennuyeux à résoudre. Par la méthode du pivot ou les formules de Cramer (par exemple), on trouve :

$$a_n = \frac{2(-4)^n + 5 \cdot 3^n - 7}{70} \quad , \quad b_n = \frac{-8(-4)^n + 15 \cdot 3^n - 7}{70}$$

$$\text{et} \quad c_n = \frac{6(-4)^n - 20 \cdot 3^n + 84}{70}.$$

On applique alors la relation (1) : $X^n = Q\chi_A + R_n$ à la matrice A : $A^n = Q(A)\chi_A(A) + R_n(A)$. D'après le théorème de Cayley-Hamilton, $\chi_A(A) = 0$ donc :

$$A^n = R_n(A) = a_n A^2 + b_n A + c_n \mathrm{I}_3 .$$

Si l'on veut vraiment avoir l'expression de la matrice A^n (particulièrement horrible ici), il suffit de remplacer. D'abord, $A^2 = \begin{pmatrix} 10 & -3 & -3 \\ -3 & 14 & 1 \\ -3 & 1 & 2 \end{pmatrix}$ puis, en remplaçant a_n, b_n, c_n par leurs valeurs, on trouve :

$$A^n = \begin{pmatrix} \frac{1}{10} + \frac{9}{14}3^n + \frac{9}{35}(-4)^n & \frac{3}{7}3^n - \frac{3}{7}(-4)^n & \frac{3}{10} - \frac{3}{14}3^n - \frac{3}{35}(-4)^n \\ \frac{3}{7}3^n - \frac{3}{7}(-4)^n & \frac{2}{7}3^n + \frac{5}{7}(-4)^n & -\frac{1}{7}3^n + \frac{1}{7}(-4)^n \\ \frac{3}{10} - \frac{3}{14}3^n - \frac{3}{35}(-4)^n & -\frac{1}{7}3^n + \frac{1}{7}(-4)^n & \frac{9}{10} + \frac{1}{14}3^n + \frac{1}{35}(-4)^n \end{pmatrix}.$$

✓ *Lorsque vous avez fini des calculs aussi lourds, il est intéressant de contrôler, au moins partiellement, la validité des résultats pour des petites valeurs de n, par exemple $n = 0$ et $n = 1$. Les formules que vous avez établies doivent donner I_3 et A respectivement, sinon...*

SF4.17 Diagonaliser la matrice

Lorsque A est diagonalisable, il existe D diagonale et P inversible telles que : $A = PDP^{-1}$. On a alors : $\forall p \in \mathbb{N}$, $A^p = PD^pP^{-1}$, formule qui se généralise à $p \in \mathbb{Z}$ dans le cas où A est inversible.
Or, si $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, on a $D^p = \mathrm{diag}(\lambda_1^p, \ldots, \lambda_n^p)$, donc le calcul de P, D, et de P^{-1} permet d'obtenir A^p.

✓ Contrairement à la méthode précédente, cette méthode exige de déterminer explicitement les sous-espaces propres, la matrice de passage et son inverse. Son avantage (le seul à notre avis) est d'être encore valable lorsque p est un entier négatif et A inversible.

Exemple

Soit $A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -2 & 1 \\ 0 & 4 & -2 \end{pmatrix}$. Calculer A^p pour $p \in \mathbb{N}$.

On peut remarquer que $A.\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$ donc 2 est valeur propre de A, et ${}^{\mathrm{t}}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ en est un vecteur propre associé.
On remarque également que les deux dernières colonnes de A sont proportionnelles ; donc A n'est pas inversible, et 0 en est une valeur propre. De plus la relation $C_2 = -2C_3$ montre que le vecteur $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ en est un vecteur propre associé.
Enfin, la trace de A étant égale à -3, la troisième valeur propre de A est égale à -5. La résolution du système $AV = -5V$ montre qu'un vecteur propre associé est le vecteur $\begin{pmatrix} 5 \\ -9 \\ 12 \end{pmatrix}$.

A possédant 3 valeurs propres distinctes est diagonalisable et $A = PDP^{-1}$ avec :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} 0 & 1 & 5 \\ 1 & 1 & -9 \\ 2 & 1 & 12 \end{pmatrix}.$$

Il nous faut donc calculer P^{-1}. Ce calcul est un peu pénible ; on obtient :

$$P^{-1} = \frac{1}{35}\begin{pmatrix} -21 & 7 & 14 \\ 30 & 10 & -5 \\ 1 & -2 & 1 \end{pmatrix}.$$

Pour tout entier naturel p on aura donc $A^p = PD^pP^{-1}$ soit, pour $p \geqslant 1$:

$$A^p = \frac{1}{35}\begin{pmatrix} 0 & 1 & 5 \\ 1 & 1 & -9 \\ 2 & 1 & 12 \end{pmatrix} \times \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2^p & 0 \\ 0 & 0 & (-5)^p \end{pmatrix} \times \begin{pmatrix} -21 & 7 & 14 \\ 30 & 10 & -5 \\ 1 & -2 & 1 \end{pmatrix}$$

et finalement, pour $p \geqslant 1$:

$$A^p = \frac{1}{35}\begin{pmatrix} 30.2^p + 5(-5)^p & 10.2^p - 10(-5)^p & -5.2^p + 5(-5)^p \\ 30.2^p - 9(-5)^p & 10.2^p + 18(-5)^p & -5.2^p - 9(-5)^p \\ 30.2^p + 12(-5)^p & 10.2^p - 24(-5)^p & -5.2^p + 12(-5)^p \end{pmatrix}.$$

✓ *Cette formule n'est pas valable pour $p = 0$ ($A^0 = \mathrm{I}_3$) ; cela vient du fait que, lors du calcul de D^p, on a écrit $0^p = 0$, ce qui n'est pas vrai pour $p = 0$.*

SF4.18 Utiliser la décomposition spectrale

Soit u un endomorphisme diagonalisable d'un espace vectoriel E. Notons $\lambda_1, \ldots, \lambda_r$ ses valeurs propres distinctes, et $E_1, \ldots, E_r$ ses sous-espaces propres associés.

À la décomposition $E = \bigoplus_{i=1}^{r} E_i$ est associée une famille de projecteurs $p_1, \ldots, p_r$: p_i est la projection sur E_i parallèlement au sous-espace vectoriel $\bigoplus_{j \neq i} E_j$.

Tout vecteur $x \in E$ s'écrit $x = \sum_{i=1}^{r} x_i$ avec pour tout i $x_i = p_i(x) \in E_i$. On a alors :

$$u(x) = \sum_{i=1}^{r} u(x_i) = \sum_{i=1}^{r} \lambda_i x_i = \sum_{i=1}^{r} \lambda_i p_i(x)$$

soit $u = \sum_{i=1}^{r} \lambda_i p_i$: c'est la *décomposition spectrale* de u.

Soit maintenant q un entier naturel ; pour tout i on a $u^q(x_i) = \lambda_i^q x_i$ (relation classique, qui se démontre par récurrence sur q). On en déduit, comme ci-dessus, que $u^q = \sum_{i=1}^{r} \lambda_i^q p_i$; ainsi, la connaissance des λ_i et des p_i permet-elle d'obtenir les puissances u^q de u.

Soit maintenant A une matrice carrée. D'après ce qui précède, lorsque A est diagonalisable, on peut écrire $A = \sum_{i=1}^{r} \lambda_i P_i$, où les P_i sont des matrices de projection.

On a alors : $\forall q \in \mathbb{N},\ A^q = \sum_{i=1}^{r} \lambda_i^q P_i$ (formule qui se généralise à $q \in \mathbb{Z}$ dans le cas où A est inversible). Il ne reste plus qu'à déterminer les P_i, ce que l'on peut faire en écrivant l'égalité précédente pour $q = 0,1,2\ldots,r-1$ (on obtient un système de r équations dont les r inconnues sont les matrices P_i, et que l'on peut résoudre par la méthode du pivot par exemple).
L'intérêt de cette méthode par rapport à la précédente est que l'on n'a pas besoin, ici, de déterminer explicitement les sous-espaces propres ni la matrice de passage. Par contre, le calcul des matrices P_i est un peu long.

✓ Dans le cas où A est diagonalisable, cette méthode est certainement plus rapide que la précédente. Son seul défaut est de ne s'appliquer, justement, qu'aux matrices diagonalisables.

Exemple

Soit $A = \begin{pmatrix} 5 & 0 & -1 \\ 1 & 4 & -1 \\ -1 & 0 & 5 \end{pmatrix}$. Calculer A^p pour $p \in \mathbb{N}$ puis pour $p \in \mathbb{Z}$.

Le calcul du polynôme caractéristique de A n'est pas détaillé ici mais nous vous conseillons de le faire à titre d'entraînement. On trouve : $\chi_A = (X-4)^2(X-6)$.

4 étant valeur propre double, pour déterminer si A est ou non diagonalisable, il nous faut connaitre la dimension du sous-espace propre associé ; or de façon claire, le rang de la matrice $A - 4\mathrm{I}_3$ est égal à 1, donc la dimension du noyau de $A - 4\mathrm{I}_3$, c'est-à-dire le sous-espace propre associé à la valeur propre 4, est égale à 2 ; A est donc diagonalisable (pour la valeur propre 6, puisqu'elle est simple, le sous-espace propre associé est forcément de dimension 1).

Il existe donc deux matrices de projections P_4 et P_6 telles que $A = 4P_4 + 6P_6$; et l'on aura aussi, pour tout $p \in \mathbb{Z}$, $A^p = 4^p P_4 + 6^p P_6$ (A est bien inversible car 0 n'est pas une valeur propre).
Il reste à déterminer P_4 et P_6 ; pour cela, on écrit les relations précédentes pour $p = 0$ et $p = 1$. Cela donne le système :

$$\begin{cases} P_4 + P_6 = \mathrm{I}_3 \\ 4P_4 + 6P_6 = A \end{cases}$$

qui se résout facilement en : $P_4 = \frac{1}{2}(6\mathrm{I}_3 - A)$ et $P_6 = \frac{1}{2}(A - 4\mathrm{I}_3)$.
On aura donc, pour tout $p \in \mathbb{Z}$:

$$A^p = \frac{1}{2}\left(4^p(6\mathrm{I}_3 - A) + 6^p(A - 4\mathrm{I}_3)\right) = (3\cdot 4^p - 2\cdot 6^p)\mathrm{I}_3 + \frac{6^p - 4^p}{2}A\,.$$

✓ A étant diagonalisable, on pouvait remarquer que le polynôme $(X-4)(X-6)$ est annulateur de A (c'est un polynôme scindé à racines simples, dont les racines sont exactement les valeurs propres de A, ou encore le polynôme minimal). La méthode de la division euclidienne donnait donc ici le résultat tout aussi rapidement, mais ne permettait pas de généraliser directement aux exposants négatifs.

SF4.19 Trigonaliser la matrice

Dans tous les cas, on peut toujours trigonaliser A, quitte à se placer dans $\mathcal{M}_n(\mathbb{C})$. Plus précisément, si A est la matrice dans une base $\mathscr{B}$ d'un certain endomorphisme u dont le polynôme caractéristique est scindé, celui-ci peut alors s'écrire sous la forme :

$$\chi_u = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}$$

où les λ_i sont les valeurs propres distinctes de u, d'ordre de multiplicité m_i.
Puisque χ_u est annulateur de u, il résulte du théorème de décomposition des noyaux que :

$$E = \bigoplus_{i=1}^{r} \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{m_i} .$$

Notons $F_i = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{m_i}$ (les F_i sont les sous-espaces caractéristiques de u). On sait que F_i est stable par u. L'endomorphisme v_i induit par u sur F_i vérifie $(v_i - \lambda_i \mathrm{Id}_{F_i})^{m_i} = 0$, c'est-à-dire que $v_i - \lambda_i \mathrm{Id}_{F_i}$ est un endomorphisme nilpotent de F_i.
Dans une base $\mathscr{B}' = \bigcup_{i=1}^{r} \mathscr{B}_i$ formée de la réunion de bases des F_i, la matrice de u sera donc diagonale par blocs, de la forme :

$$T = \begin{bmatrix} \boxed{T_1} & & 0 \\ & \boxed{T_2} & \\ & & \ddots \\ 0 & & \boxed{T_r} \end{bmatrix}$$

où chaque T_i est triangulaire supérieure, de la forme $T_i = \lambda_i \mathrm{I}_{m_i} + N_i$ avec N_i nilpotente (un exemple de calcul avec cette méthode est donné dans l'exercice **4.1**).
On sait alors calculer les T_i^p pour $p \in \mathbb{N}$ par la formule du binôme, et on a ensuite :

$$T^p = \begin{bmatrix} \boxed{T_1^p} & & 0 \\ & \boxed{T_2^p} & \\ & & \ddots \\ 0 & & \boxed{T_r^p} \end{bmatrix} .$$

Si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, on a $A = PTP^{-1}$ puis $A^p = PT^pP^{-1}$.

✓ Cette méthode donne lieu en général à des calculs pénibles.

Exemple

Soit $A = \begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

1) Montrer que A est semblable à la matrice $B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

2) Calculer A^n pour tout $n \in \mathbb{N}$.

1) *Remarquons déjà que, puisque l'énoncé donne le résultat à obtenir, on n'a pas ici besoin de calculer le polynôme caractéristique de A.*
Si A est la matrice dans la base canonique (e_1,e_2,e_3) de $\mathbb{R}^3$ d'un certain endomorphisme $u \in \mathscr{L}(\mathbb{R}^3)$, il s'agit de trouver une autre base (e'_1,e'_2,e'_3) de $\mathbb{R}^3$ telle que la matrice de u dans cette base soit égale à B.
Cela équivaut aux relations :

$$u(e'_1) = 3e'_1 \quad , \quad u(e'_2) = e'_2 \quad \text{et} \quad u(e'_3) = e'_2 + e'_3 \, .$$

✓ *Remarquez que les deux premières relations reviennent à trouver des vecteurs propres pour les valeurs propres 3 et 1.*

En notant (x,y,z) les coordonnées de e'_1 dans la base initiale, on a :

$$u(e'_1) = 3e'_1 \iff \left\{ \begin{array}{rl} 3x + z & = 3x \\ 2x + y + z & = 3y \\ -x + y + z & = 3z \end{array} \right. \iff \left\{ \begin{array}{l} x = y \\ z = 0 \end{array} \right. .$$

On peut donc choisir e'_1 de coordonnées $(1,1,0)$.
Puis en notant (x,y,z) les coordonnées de e'_2 dans la base initiale, on a :

$$u(e'_2) = e'_2 \iff \left\{ \begin{array}{rl} 3x + z & = x \\ 2x + y + z & = y \\ -x + y + z & = z \end{array} \right. \iff \left\{ \begin{array}{l} x = \ \ y \\ z = -2x \end{array} \right. .$$

On peut donc choisir e'_2 de coordonnées $(1,1,-2)$.
Enfin, en notant (x,y,z) les coordonnées de e'_3 dans la base initiale, on a :

$$u(e'_3) = e'_3 + e'_2 \iff \left\{ \begin{array}{rl} 3x + z & = x + 1 \\ 2x + y + z & = y + 1 \\ -x + y + z & = z - 2 \end{array} \right. \iff \left\{ \begin{array}{l} 2x + z = 2 \\ -x + y = 1 \end{array} \right. .$$

On peut choisir e'_3 de coordonnées $(1,2,0)$.
La matrice des vecteurs e'_1, e'_2, e'_3 dans la base canonique est :

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & -2 & 0 \end{pmatrix} .$$

Cette matrice est bien inversible, donc (e'_1,e'_2,e'_3) est bien une base de $\mathbb{R}^3$ dans laquelle la matrice de u est B.

2) *Les formules habituelles de changement de base donnent alors $A = PBP^{-1}$, d'où $A^n = PB^nP^{-1}$ pour tout $n \in \mathbb{N}$.*
La matrice B peut s'écrire par blocs $B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & & \\ 0 & & C \end{bmatrix}$ avec $C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ soit $C = \mathrm{I}_2 + N$, et $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ nilpotente ($N^2 = 0_2$).
Par la formule du binôme, on a $C^n = \mathrm{I}_2 + nN = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. On en déduit $B^n = \begin{pmatrix} 3^n & 0 & 0 \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$. Pour terminer le calcul de A^n, il ne reste « plus qu'à » calculer P^{-1}, puis effectuer le produit PA^nP^{-1}.

✓ *C'est volontairement (et un peu par paresse !) que nous n'avons pas fini le calcul : en effet, une fois obtenu le polynôme caractéristique, la méthode de la division euclidienne est, encore une fois, bien plus rapide !*

■ Comment utiliser la réduction dans d'autres problèmes ?

Nous donnons dans cette partie quelques exemples où il faut penser à utiliser la réduction, même si cela n'est pas toujours précisé par l'énoncé. Il faut apprendre à reconnaître ce genre de situations.

SF4.20 Étudier des systèmes linéaires de suites récurrentes

Exemple

Déterminer les suites (u_n), (v_n) et (w_n) qui vérifient $(u_0,v_0,w_0) = (1,0,0)$ et les relations de récurrence valables pour tout $n \in \mathbb{N}$:

$$\begin{cases} u_{n+1} = -u_n + v_n + w_n \\ v_{n+1} = u_n - v_n + w_n \\ w_{n+1} = u_n + v_n - w_n \end{cases}.$$

- *Pour résoudre ce genre d'exercices, il suffit de remarquer que, si l'on pose*

$$X_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} \quad \text{et} \quad A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix},$$

le système de l'énoncé est équivalent à : $X_{n+1} = AX_n$.

On a alors, par récurrence immédiate : $\forall\, n \in \mathbb{N},\ X_n = A^n X_0$.

- *On est donc ramené à calculer* A^n. *Or* A *est une matrice symétrique réelle, donc diagonalisable ; de plus la somme de chaque ligne est égale à* 1 *donc* 1 *est valeur propre ; puisque* $\operatorname{tr} A = -3$ *et* $\det A = 4$, *on trouve que les deux autres valeurs propres sont* -2 *et* -2.
A étant diagonalisable, on sait d'après **[S4.10]** *que le polynôme* $(X-1)(X+2)$ *est annulateur de* A. *La division euclidienne de* X^n *par ce polynôme s'écrit :*

$$X^n = (X-1)(X+2)Q + a_n X + b_n\,,$$

et en appliquant cette relation en 1 *puis* -2 *:*

$$1 = a_n + b_n \quad \text{et} \quad (-2)^n = -2a_n + b_n\,,$$

soit

$$a_n = \frac{1-(-2)^n}{3} \quad \text{et} \quad b_n = \frac{2+(-2)^n}{3}.$$

On obtient finalement :

$$A^n = a_n A + b_n \mathrm{I}_3 = \frac{1}{3}\begin{pmatrix} 1+2(-2)^n & 1-(-2)^n & 1-(-2)^n \\ 1-(-2)^n & 1+2(-2)^n & 1-(-2)^n \\ 1-(-2)^n & 1-(-2)^n & 1+2(-2)^n \end{pmatrix}.$$

Pour terminer, on aura $\begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = A^n . \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ *ce qui donne :*

$$u_n = \frac{1+2(-2)^n}{3},\ v_n = \frac{1-(-2)^n}{3} \quad \text{et} \quad w_n = \frac{1-(-2)^n}{3}.$$

✓ *Étant donné les conditions initiales* $(1,0,0)$, *seule la 1re colonne de* A^n *était utile pour conclure ; on n'avait donc pas besoin de calculer les deux autres.*

SF4.21 Étudier des suites récurrentes linéaires

Nous avons vu en **[S8.6]** comment étudier les suites récurrentes linéaires d'ordre 2. Nous nous proposons ici de généraliser les résultats obtenus.

Soit $E = \mathbb{C}^{\mathbb{N}}$ le $\mathbb{C}$-espace vectoriel des suites complexes. Étant donnés un entier $p \geqslant 2$ et $(a_0, \dots, a_{p-1}) \in \mathbb{C}^p$, on veut étudier l'ensemble des suites vérifiant une relation de récurrence linéaire d'ordre p :

$$\mathscr{S} = \{u \in E \mid \forall\, n \in \mathbb{N},\ u_{n+p} = a_{p-1}u_{n+p-1} + \cdots + a_1 u_{n+1} + a_0 u_n\}\ .$$

✓ Nous ne ferons pas ici l'étude théorique du problème ; nous nous contentons d'exposer la méthode pratique pour trouver les éléments de $\mathscr{S}$ en utilisant le calcul matriciel.

Soit $(u_n) \in \mathscr{S}$ et $X_n = \begin{pmatrix} u_n \\ \vdots \\ u_{n+p-1} \end{pmatrix}$. On suppose que $u_0, u_1, \dots, u_{p-1}$ sont donnés.

On connaît donc X_0, et on a $X_{n+1} = AX_n$ où A est la matrice :

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ a_0 & a_1 & \dots & \dots & a_{p-1} \end{pmatrix}.$$

On a alors pour tout entier n : $X_n = A^n X_0$ (récurrence immédiate), et le calcul de A^n permet d'obtenir u_n en fonction de n.

Exemple

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$u_0 = 1\,,\ u_1 = 0\,,\ u_2 = 11\,,\ u_3 = 20\,,$$

et la relation de récurrence :

$$\forall\, n \in \mathbb{N},\ u_{n+4} = 2u_{n+3} + 3u_{n+2} - 4u_{n+1} - 4u_n\,.$$

Calculer u_n en fonction de n.

En notant $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \\ u_{n+3} \end{pmatrix}$ *les données précédentes s'écrivent :*

$$X_0 = \begin{pmatrix} 1 \\ 0 \\ 11 \\ 20 \end{pmatrix} \text{ et } X_{n+1} = AX_n \text{ avec } A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -4 & -4 & 3 & 2 \end{pmatrix}.$$

On est donc ramené au calcul de A^n. *Le polynôme caractéristique de* A *est*

$$\chi_A = X^4 - 2X^3 - 3X^2 + 4X + 4 = (X+1)^2(X-2)^2\,.$$

Pour calculer A^n, la méthode de la division euclidienne est la plus rapide (puisque l'on n'a pas besoin ici des éléments propres).
La division euclidienne de X^n par χ_A s'écrit :

$$X^n = (X+1)^2(X-2)^2 + aX^3 + bX^2 + cX + d\,.$$

En évaluant cette expression ainsi que sa dérivée en -1 puis en 2, on obtient le système :

$$\begin{cases} (-1)^n &= -a+b-c+d \\ n(-1)^n &= 3a-2b+c \\ 2^n &= 8a+4b+2c+d \\ n2^n &= 12a+4b+c \end{cases}.$$

Puisque χ_A est annulateur de A d'après le théorème de Cayley-Hamilton, on a :

$$A^n = aA^3 + bA^2 + cA + d\mathrm{I}_4$$

puis u_n sera la première ligne du produit matriciel A^nX_0.
L'astuce consiste ici à ne pas faire les calculs mais à deviner ce qui se passerait si on les faisait ! Nous vous laissons le soin de réfléchir pour être convaincu que l'on obtiendra forcément une expression de la forme :

$$u_n = \alpha(-1)^n + \beta n(-1)^n + \gamma 2^n + \delta n2^n\,.$$

Il ne reste plus qu'à déterminer α, β, γ et δ en considérant les cas $n = 0$, 1, 2, 3. On trouve finalement $\alpha = \beta = \delta = 1$ et $\gamma = 0$ donc

$$\forall\, n \in \mathbb{N},\ u_n = (n+1)(-1)^n + n2^n\,.$$

SF4.22 Étudier le commutant d'un endomorphisme ou d'une matrice

Soit E un $\mathbb{K}$-espace vectoriel, et u un endomorphisme de E. On appelle commutant de u l'ensemble $\mathscr{C}(u)$ des endomorphismes de E qui commutent avec u :

$$\mathscr{C}(u) = \{v \in \mathscr{L}(E) \mid u \circ v = v \circ u\}\,.$$

Il est facile de vérifier que $\mathscr{C}(u)$ est une sous-algèbre de $\mathscr{L}(E)$.
On définit de manière similaire le commutant d'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$:

$$\mathscr{C}(A) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid AM = MA\}\,.$$

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{K})$, diagonalisable. Donner la dimension du commutant de A :

$$\mathscr{C}(A) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid AM = MA\}$$

en fonction des ordres de multiplicité des valeurs propres de A.

Notons u l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A, $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de u, $E_{\lambda_i}(u)$ les sous-espaces propres de u, et $\mathscr{B}_i$ une base de $E_{\lambda_i}(u)$.
L'énoncé équivaut à la recherche de la dimension du commutant $\mathscr{C}(u)$.

✓ *Comme souvent, il est plus avantageux d'utiliser les applications linéaires que de traiter le problème seulement du point de vue matriciel.*

u étant diagonalisable, la famille $\mathscr{B} = \bigcup_{i=1}^{p} \mathscr{B}_i$ est une base de E dans laquelle la matrice U de u est de la forme :

$$U = \begin{pmatrix} \boxed{D_1} & & & 0 \\ & \boxed{D_2} & & \\ & & \ddots & \\ 0 & & & \boxed{D_p} \end{pmatrix},$$

chaque bloc D_i étant une matrice scalaire $\mathrm{diag}(\lambda_i,\ldots,\lambda_i)$ *de taille m_i où m_i est l'ordre de multiplicité de la valeur propre λ_i.*

Soit $v \in \mathscr{C}(u)$. On a vu en **[S4.6]** *que v laisse stable les $E_{\lambda_i}(u)$. Cela implique (cf.* **[S3.8]***) que la matrice de v dans la base $\mathscr{B}$ est de la forme :*

$$V = \begin{pmatrix} \boxed{V_1} & & & 0 \\ & \boxed{V_2} & & \\ & & \ddots & \\ 0 & & & \boxed{V_p} \end{pmatrix} \qquad (*)$$

chaque bloc V_i étant une matrice carrée quelconque de taille m_i.

Réciproquement, si la matrice V est de la forme ci-dessus, elle commute avec U puisque chaque bloc V_i commute avec la matrice scalaire D_i.

La dimension cherchée est donc égale, par isomorphisme, à celle de l'espace vectoriel des matrices de la forme $(*)$ c'est-à-dire à $\sum_{i=1}^{p} m_i^2$.

SF4.23 Résoudre des équations matricielles

L'idée pour résoudre une telle équation est de diagonaliser ou de trigonaliser les matrices données, afin de se ramener au même type de problème mais avec des matrices plus simples.

Il est très fréquent d'utiliser dans ces problèmes le résultat important suivant, déjà rappelé en **[S3.5]** et démontré en **[SF3.1]**.

> Si une matrice carrée commute avec une matrice diagonale à éléments diagonaux distincts, elle est diagonale.

Exemple

Trouver les matrices $M \in \mathcal{M}_3(\mathbb{R})$ telles que $M^3 = A$, avec $A = \begin{pmatrix} -4 & 4 & -4 \\ -4 & 4 & 4 \\ -8 & 8 & 0 \end{pmatrix}$.

Traiter la même question dans $\mathcal{M}_3(\mathbb{C})$.

- *On commence par diagonaliser A.*

  ✓ *Nous avons déjà traité un certain nombre d'exemples de diagonalisation de matrices 3×3. C'est maintenant à vous de vous exercer ! Nous ne donnons ici que les résultats, sans détailler les calculs.*

  On trouve : $\chi_A = X^3 - 64X$, puis $A = PDP^{-1}$ avec $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ et $D = \mathrm{diag}\,(0,8,-8)$. On calcule également $P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$.

- *Soit $M \in \mathcal{M}_3(\mathbb{R})$ telle que $M^3 = A$; posons $M' = P^{-1}MP$. Alors :*

  $M^3 = A \iff (PM'P^{-1})^3 = A \iff PM'^3P^{-1} = PDP^{-1} \iff M'^3 = D$.

  M' commute avec D puisque $DM' = M'^3M' = M'^4 = M'D$.
  D'après la remarque faite ci-dessus, D étant à éléments diagonaux distincts, on en conclut que M' est diagonale.

  Ainsi $M' = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ *avec a, b, c réels ; la relation $M'^3 = D$ équivaut alors à* $\begin{cases} a^3 = 0 \\ b^3 = 8 \\ c^3 = -8 \end{cases}$ *d'où* $\begin{cases} a = 0 \\ b = 2 \\ c = -2 \end{cases}$, *et enfin* $M' = \operatorname{diag}(0,2,-2)$.

  Finalement $M = P \cdot \operatorname{diag}(0,2,-2) \cdot P^{-1} = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & 1 \\ -2 & 2 & 0 \end{pmatrix}$.

- *Lorsqu'on résout l'équation dans $\mathcal{M}_3(\mathbb{C})$, la seule différence est dans la résolution du dernier système, qui donne les solutions $a = 0$, $b \in \{2, 2\mathrm{j}, 2\bar{\mathrm{j}}\}$ et $c \in \{-2, -2\mathrm{j}, -2\bar{\mathrm{j}}\}$, avec $\mathrm{j} = \mathrm{e}^{2\mathrm{i}\pi/3}$, soit au final 9 solutions pour M.*

Pour résoudre une équation matricielle dans $\mathcal{M}_n(\mathbb{R})$, on peut être amené à se placer provisoirement dans $\mathcal{M}_n(\mathbb{C})$, et parfois à utiliser le résultat démontré en **[SF3.30]** page 273 :

> Si deux matrices à coefficients réels sont semblables dans $\mathcal{M}_n(\mathbb{C})$, elles sont aussi semblables dans $\mathcal{M}_n(\mathbb{R})$.

Exemple

Caractériser toutes les matrices $A \in \mathcal{M}_n(\mathbb{R})$ telles que

$$A^2 + A + \mathrm{I}_n = 0_n .$$

- *Si λ est une valeur propre de A, on a $\lambda^2 + \lambda + 1 = 0$, donc $\lambda \in \{\mathrm{j}, \mathrm{j}^2\}$. La matrice A n'ayant pas de valeur propre réelle, on en déduit que n est pair (puisque tout polynôme de $\mathbb{R}[X]$ de degré impair admet au moins une racine).*

- *Posons donc $n = 2p$. Puisque A annule un polynôme scindé à racines simples dans $\mathbb{C}[X]$, elle est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$.*
  Le polynôme caractéristique de A étant à coefficients réels, ses racines complexes sont deux à deux conjuguées et les valeurs propres de A ont même ordre de multiplicité p ; de plus, si $(V_1, \ldots, V_p)$ est une base de $E_{\mathrm{j}}(A)$, alors $(\overline{V_1}, \ldots, \overline{V_p})$ est une base de $E_{\mathrm{j}^2}(A)$.
  On a alors, pour tout $k \in [\![1\,;p]\!]$:

$$A(V_k + \overline{V_k}) = \mathrm{j}V_k + \mathrm{j}^2\overline{V_k}$$

$$\text{et} \quad A(\mathrm{j}V_k + \mathrm{j}^2\overline{V_k}) = \mathrm{j}^2V_k + \mathrm{j}\overline{V_k} = (-1-\mathrm{j})V_k + (-1-\mathrm{j}^2)\overline{V_k} ,$$

  de sorte qu'en considérant les vecteurs à coordonnées réelles :

$$W_k = V_k + \overline{V_k} \quad \text{et} \quad W'_k = \mathrm{j}V_k + \mathrm{j}^2\overline{V_k}$$

  on a, pour tout $k \in [\![1\,;p]\!]$:

$$AW_k = W'_k \quad \text{et} \quad AW'_k = -W_k - W'_k .$$

- *Enfin il est facile de vérifier que la famille* $(W_1,W_1',\dots,W_p,W_p')$ *est libre dans* $\mathcal{M}_{2p,1}(\mathbb{R})$*, donc en forme une base. Dans cette base, la matrice de l'endomorphisme canoniquement associé à* A *sera égale à la matrice par blocs* $\operatorname{diag}(\underbrace{J,J,\dots,J}_{p \text{ blocs}})$ *où* $J=\begin{pmatrix}0&-1\\1&-1\end{pmatrix}$*, c'est-à-dire que* A *est semblable à cette matrice.*
  Réciproquement, puisque $J^2+J+\mathrm{I}_2=0_2$*, toute matrice semblable à* $\operatorname{diag}(J,J,\dots,J)$ *vérifie l'équation proposée.*
- *Autre solution possible : en considérant la base* $(V_1,\overline{V_1},\dots,V_p,\overline{V_p})$*, on remarque que* A *est semblable dans* $\mathcal{M}_n(\mathbb{C})$ *à la matrice* $\operatorname{diag}(K,K,\dots,K)$ *où* $K=\begin{pmatrix}\mathrm{j}&0\\0&\mathrm{j}^2\end{pmatrix}$*. Puisque* K *et* J *sont semblables (car* J *diagonalisable de valeurs propres* j *et* j^2*), on en déduit que* A *et* $\operatorname{diag}(J,J,\dots,J)$ *sont semblables dans* $\mathcal{M}_n(\mathbb{C})$ *donc aussi dans* $\mathcal{M}_n(\mathbb{R})$*, en vertu de la remarque faite plus haut.*

SF4.24 Étudier l'inversibilité d'une matrice

Pour savoir si une matrice carrée est inversible, on peut rechercher ses valeurs propres. Si elles sont toutes non nulles, la matrice est inversible.

Exemple

Soient a ,$b\in\mathbb{C}$. Étudier selon les valeurs de a et b l'inversibilité de la matrice :

$$A=\begin{pmatrix}a&b&0&\cdots&0&b\\b&a&b&\cdots&0&0\\0&b&a&\ddots&0&0\\\vdots&\ddots&\ddots&\ddots&\ddots&\vdots\\0&0&0&\ddots&a&b\\b&0&0&\cdots&b&a\end{pmatrix}\in\mathcal{M}_n(\mathbb{C})(n\geqslant 2)\,.$$

On commence par remarquer que la matrice A *s'écrit :* $A=a\mathrm{I}_n+bK$ *où* K *est la matrice obtenue pour* $a=0$ *et* $b=1$*. Pour trouver les valeurs propres de* A*, il suffit donc de trouver celles de* K *(cf.* **[SF4.8]***).*

Or :

$$K=\begin{pmatrix}0&1&0&\cdots&0&1\\1&0&1&\cdots&0&0\\0&1&0&\ddots&0&0\\\vdots&\ddots&\ddots&\ddots&\ddots&\vdots\\0&0&0&\ddots&0&1\\1&0&0&\cdots&1&0\end{pmatrix},$$

et si l'on pose :

$$J = \begin{pmatrix} 0 & 1 & 0 & & & \\ \vdots & \ddots & \ddots & \ddots & & 0 \\ \vdots & & \ddots & \ddots & 0 & \\ \vdots & & 0 & \ddots & 1 & 0 \\ 0 & & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

on a $K = J + {}^{\mathrm{t}}J = J + J^{-1}$ *(la matrice* J *est orthogonale, ses vecteurs colonnes formant une base orthonormale de* $\mathbb{R}^n$*).*

Or on a vu dans l'exemple traité page 300 que la matrice J *est diagonalisable :* $J = PDP^{-1}$ *où* $D = \mathrm{diag}(1,\omega,\ldots,\omega^{n-1})$ *avec* P *inversible et* $\omega = \mathrm{e}^{\frac{2\mathrm{i}\pi}{n}}$*. On aura donc* $J^{-1} = PD^{-1}P^{-1}$ *puis* $K = P(D + D^{-1})P^{-1}$*.*

Finalement les valeurs propres de K *sont les réels :*

$$\omega_k + \omega_k^{-1} = \omega_k + \overline{\omega_k} = 2\mathcal{R}e\,(\omega_k)$$

avec $\omega_k = \mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}$ *et* $k \in [\![0\,;n-1]\!]$*, c'est-à-dire les réels :*

$$2\cos\left(\frac{2k\pi}{n}\right) \quad k \in \left[\!\!\left[0\,;\left\lfloor\frac{n}{2}\right\rfloor\right]\!\!\right],$$

les valeurs propres autres que 2 *et* -2 *(dans le cas* n *pair) étant doubles.*

Il en résulte que les valeurs propres de la matrice A *donnée dans l'énoncé sont les nombres* $a+2b\cos\left(\frac{2k\pi}{n}\right)$ *avec* $k \in \left[\!\!\left[0\,;\left\lfloor\frac{n}{2}\right\rfloor\right]\!\!\right]$ *et donc que* A *est inversible si et seulement si il n'existe pas d'entier* k *tel que* $a+2b\cos\left(\frac{2k\pi}{n}\right) = 0$*.*

✓ Cet exemple montre qu'il est important de bien connaître un certains exemples classiques, ainsi que certaines matrices carrées remarquables.

SF4.25 Étudier des polynômes de matrice

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $P \in \mathbb{C}[X]$. Montrer que la matrice $P(A)$ est inversible si et seulement si P est premier avec le polynôme minimal π_A de la matrice A.

- *Supposons les polynômes* P *et* π_A *premiers entre eux.*
  D'après le théorème de Bezout, il existe des polynômes U *et* V *tels que :*

  $$U\pi_A + VP = 1\,,$$

  ce qui implique $U(A)\pi_A(A) + V(A)P(A) = \mathrm{I}_n$ *puis* $V(A)P(A) = \mathrm{I}_n$ *puisque* π_A *est annulateur de* A*.*
  Donc $P(A)$ *est inversible, d'inverse* $V(A)$*.*
- *Réciproquement, supposons les polynômes* P *et* π_A *non premiers entre eux, et notons* D *leur pgcd.* D *est donc un polynôme non constant. On peut écrire* $P \cdot \left(\frac{\pi_A}{D}\right) = \pi_A \cdot \left(\frac{P}{D}\right)$*, donc il existe un polynôme* Q *tel que* $\deg Q < \deg \pi_A$ *et un polynôme* R *tels que* $PQ = \pi_A R$*.*

On a donc $P(A)Q(A) = \pi_A(A)R(A) = 0_n$; or $Q(A) \neq 0_n$ par définition du polynôme minimal, donc la matrice $P(A)$ n'est pas inversible.
Par contraposition, on a donc prouvé que $P(A)$ inversible implique P et π_A premiers entre eux.

SF4.26 Calculer trace et déterminant d'un endomorphisme

Il s'agit ici d'utiliser la propriété suivante : si le polynôme caractéristique d'un endomorphisme u est scindé, la trace de u égale à la somme des valeurs propres et le déterminant de u à leur produit.

Exemple

Soit u l'endomorphisme de $\mathcal{M}_n(\mathbb{K})$ défini par $u\colon M \mapsto a\,{}^tM + bM$, où a et b sont des scalaires donnés.

Déterminer la trace et le déterminant de u.

Cherchons les éléments propres de u, en résolvant l'équation $u(M) = \lambda M$.
Pour cela, on remarque que si M est une matrice symétrique ($M \in \mathcal{S}_n(\mathbb{K})$) on a $u(M) = (a+b)M$, et que si M est une matrice antisymétrique ($M \in \mathcal{A}_n(\mathbb{K})$) on a $u(M) = (b-a)M$.
Ainsi $a+b$ est une valeur propre de u, le sous-espace propre associé contenant $\mathcal{S}_n(\mathbb{K})$ et $b-a$ est une valeur propre de u, le sous-espace propre associé contenant $\mathcal{A}_n(\mathbb{K})$.

- *Si $a = 0$, ces deux valeurs propres sont égales ; u est alors l'homothétie de rapport b.*
- *Sinon, les deux valeurs propres sont distinctes, et puisque les sous-espaces vectoriels $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$, on en déduit que u est diagonalisable. Il en résulte que $a+b$ est d'ordre $\dim \mathcal{S}_n(\mathbb{K}) = \dfrac{n(n+1)}{2}$ et que $b-a$ est d'ordre $\dim \mathcal{A}_n(\mathbb{K}) = \dfrac{n(n-1)}{2}$.*

On en conclut que :

$$\operatorname{tr} u = \frac{n(n+1)}{2}(a+b) + \frac{n(n-1)}{2}(b-a) = na + n^2 b$$

et que :

$$\det u = (a+b)^{\frac{n(n+1)}{2}}(b-a)^{\frac{n(n-1)}{2}},$$

ces formules restant vraies dans le cas $a = 0$.

SF4.27 Obtenir certains renseignements sur une matrice

Exemple

Soit $A \in \mathcal{M}_5(\mathbb{R})$ telle que $12A^3 - 8A^2 + 7A - I_5 = 0$. Montrer que $0 < \operatorname{tr} A < 2$.

Le polynôme $P = 12X^3 - 8X^2 + 7X - 1$, annulateur de A, n'a qu'une seule racine réelle, qui est $\dfrac{1}{6}$. Il possède deux autres racines complexes conjuguées λ et $\overline{\lambda}$. On sait aussi (relations coefficients-racines pour un polynôme) que $\dfrac{1}{6} + \lambda + \overline{\lambda} = \dfrac{2}{3}$, soit $\lambda + \overline{\lambda} = \dfrac{1}{2}$.

P étant scindé à racines simples dans $\mathbb{C}[X]$, on peut diagonaliser A dans $\mathcal{M}_5(\mathbb{C})$. De plus, $\frac{1}{6}$ est forcément valeur propre de A (car le polynôme caractéristique de A est à coefficients réels, de degré impair, donc possède au moins une racine réelle), et les ordres de multiplicité de λ et $\overline{\lambda}$ sont identiques.

Donc les valeurs propres de A ne peuvent être que $\left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right)$ ou $\left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\lambda,\overline{\lambda}\right)$ ou $\left(\frac{1}{6},\lambda,\lambda,\overline{\lambda},\overline{\lambda}\right)$.

La trace de A vaut alors suivant les cas $\frac{5}{6}$, $\frac{3}{6}+\lambda+\overline{\lambda}=1$ ou $\frac{1}{6}+2(\lambda+\overline{\lambda})=\frac{7}{6}$. Dans les trois cas, on a bien l'encadrement demandé.

SF4.28 Rechercher des sous-espaces stables

Il est facile de trouver les sous-espaces vectoriels stables par un endomorphisme diagonalisable : ce sont exactement ceux dont une base est formée de vecteurs propres.
Dans le cas d'un endomorphisme qui annule un polynôme scindé, nous avons cité en **[S4.14]** un théorème qui caractérise les sous-espaces stables, mais ce théorème n'est guère utilisable en pratique.
On peut cependant penser à utiliser l'un des résultats suivants.

- Si F est stable par u, alors le polynôme caractéristique de l'endomorphisme induit par u sur F divise le polynôme caractéristique de u.
- Dire qu'une droite $D = \text{Vect}(x)$ est stable par u équivaut à dire que x est un vecteur propre de u.

Exemple

1) Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, rapporté à une base $\mathscr{B}$, et soit H l'hyperplan d'équation :

$$a_1x_1 + \cdots + a_nx_n = 0 \ \left((a_1,\ldots,a_n) \neq (0,\ldots,0)\right).$$

Soit u un endomorphisme de E, et A sa matrice dans $\mathscr{B}$.

Montrer que : H stable par $u \iff {}^{\mathrm{t}}\begin{pmatrix}a_1 & \ldots & a_n\end{pmatrix}$ vecteur propre de ${}^{\mathrm{t}}A$.

2) Déterminer les sous-espaces vectoriels de $\mathbb{R}^3$ stables par l'endomorphisme u dont la matrice dans la base canonique de $\mathbb{R}^3$ est $A = \begin{pmatrix}1 & 2 & -1\\ 2 & 1 & 1\\ 0 & 0 & -1\end{pmatrix}$.

1) *La base $\mathscr{B}$ étant donnée, on pourra confondre tout vecteur x de E avec la matrice colonne X de ses coordonnées dans $\mathscr{B}$.*
Notons L la matrice ligne $\begin{pmatrix}a_1 & a_2 & \ldots & a_n\end{pmatrix}$. Ainsi, la relation $x \in H$ peut s'écrire matriciellement $LX = 0$.

- *Supposons H stable par u. Cela signifie que l'on a l'implication :*

$$\forall\, x \in E,\ x \in H \implies u(x) \in H$$

soit :

$$\forall\, X \in \mathcal{M}_{n,1}(\mathbb{K}),\ LX = 0 \implies LAX = 0\,.$$

Or l'application $\varphi : X \mapsto LX$ et l'application $\psi : X \mapsto LAX$ sont des formes linéaires sur $\mathcal{M}_{n,1}(\mathbb{K})$ (dont les matrices dans les bases canoniques sont précisément L et LA). L'implication ci-dessus peut aussi s'écrire :

$$X \in \operatorname{Ker}\varphi \Longrightarrow X \in \operatorname{Ker}\psi$$

soit : $\operatorname{Ker}\varphi \subset \operatorname{Ker}\psi$. φ étant non nulle, cela implique qu'il existe un scalaire λ tel que $\psi = \lambda\varphi$ (cf. **[S2.20]**).
On a donc :

$$\forall\, X \in \mathcal{M}_{n,1}(\mathbb{K}),\ LAX = \lambda LX$$

d'où $LA = \lambda L$ (car ce sont les matrices de la même forme linéaire), puis ${}^tA\,{}^tL = \lambda\,{}^tL$.
Cela montre que ${}^tL$ est un vecteur propre de ${}^tA$ (il est bien non nul).

- *Réciproquement, si ${}^tL$ est un vecteur propre de ${}^tA$, il existe λ tel que ${}^tA\,{}^tL = \lambda\,{}^tL$ et en remontant le raisonnement précédent, on obtient que H est stable par u.*

2) *On va rechercher les sous-espaces stables par u en les classant selon leur dimension.*

- *$\{0\}$ et $\mathbb{R}^3$ sont des sous-espaces vectoriels stables par n'importe quel endomorphisme.*
- *Pour trouver les droites stables par u on est amené à chercher ses vecteurs propres. Un calcul rapide montre que $\operatorname{Sp} A = \{-1,3\}$, que le sous-espace propre associé à la valeur propre -1 est la droite de base (a) avec $a = (1,-1,0)$ et que le sous-espace propre associé à la valeur propre 3 est la droite de base (b) avec $b = (1,1,0)$.*
  Ces sous-espaces n'étant pas supplémentaires, A n'est pas diagonalisable ; et les seules droites stables par u sont les deux droites précédentes.
- *Pour trouver les plans stables par u, on cherche les sous-espaces propres de ${}^tA$. Les valeurs propres de ${}^tA$ sont les mêmes que celles de A ; on trouve que le sous-espace propre associé à la valeur propre -1 est la droite de base le vecteur $(0,0,1)$ et que le sous-espace propre associé à la valeur propre 3 est la droite de base le vecteur $(1,1,0)$.*
  D'après la question **1)**, *les plans stables par u sont donc ceux d'équation $z = 0$ et $x + y = 0$.*
  On peut remarquer que le premier de ces plans est le plan $\operatorname{Vect}(a,b)$ et que le second contient le vecteur a.

SF4.29 Calculer l'exponentielle d'une matrice

Rappelons (voir **[S11.9]**) que pour toute matrice carrée $A \in \mathcal{M}_p(\mathbb{K})$, on peut définir l'exponentielle de A par : $\exp(A) = \displaystyle\sum_{n=0}^{+\infty} \frac{1}{n!} A^n$.

La principale propriété utilisée pour le calcul de cette matrice est la suivante.

Si $B = P^{-1}AP$ avec $P \in \mathrm{GL}_p(\mathbb{K})$, alors $\exp(B) = P^{-1}\exp(A)P$.

En effet, on a, pour tout entier $k \in \mathbb{N}$, $B^k = P^{-1}A^kP$, donc

$$\forall\, n \in \mathbb{N},\ \sum_{k=0}^{n} \frac{1}{k!} B^k = P^{-1} \left(\sum_{k=0}^{n} \frac{1}{k!} A^k \right) P\,.$$

Le résultat annoncé s'obtient alors par passage à la limite quand $n \to +\infty$, ce qui est justifié car l'application $M \mapsto P^{-1}MP$ de $\mathcal{M}_n(\mathbb{K})$ dans $\mathcal{M}_n(\mathbb{K})$ est continue (car c'est un endomorphisme d'un espace vectoriel de dimension finie).

Pour calculer l'exponentielle d'une matrice, il « suffit » donc de la diagonaliser ou de la trigonaliser. En effet :

- si $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, on a facilement $\exp(D) = \operatorname{diag}(\mathrm{e}^{\lambda_1}, \ldots, \mathrm{e}^{\lambda_n})$;
- si T est une matrice triangulaire de la forme $T = \lambda \mathrm{I} + N$ avec N nilpotente, on aura $\exp(T) = \exp(\lambda \mathrm{I}) \exp(N) = \mathrm{e}^{\lambda} \exp(N)$, avec $\exp(N) = \displaystyle\sum_{k=0}^{r-1} \frac{1}{k!} N^k$, où r est l'indice de nilpotence de N ;
- enfin, si T est une matrice diagonale par blocs avec $T = \operatorname{diag}(T_1, \ldots, T_s)$ où les T_i sont des matrices triangulaires supérieures de la forme précédente, la matrice $\exp(T)$ est la matrice diagonale par blocs $\exp(T) = \operatorname{diag}\big(\exp(T_1), \ldots, \exp(T_s)\big)$.

Exemple

Soit $a \in \mathbb{R}^*$. Calculer l'exponentielle de la matrice $A = \begin{pmatrix} -1 & a & a \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$.

Nous vous laissons le soin de trigonaliser cette matrice par les méthodes habituelles. On trouve (par exemple) :

$$A = PTP^{-1} \quad \textit{avec} \quad T = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad \textit{et} \quad P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \frac{1}{a} \\ -1 & 0 & 0 \end{pmatrix}.$$

En écrivant $T = -\mathrm{I}_3 + N$ avec $N^3 = 0_3$, on a :

$$\exp(N) = \mathrm{I}_3 + N + \frac{1}{2}N^2 = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \textit{puis} \quad \exp(T) = \mathrm{e}^{-1} \exp(N)\,.$$

On en déduit finalement, après calcul de P^{-1} :

$$\exp(A) = P \exp(T) P^{-1} = \mathrm{e}^{-1} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \frac{1}{a} \\ -1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & a & a \end{pmatrix}$$

$$= \mathrm{e}^{-1} \begin{pmatrix} 1 & a & a \\ 1 & \frac{a}{2} + 1 & \frac{a}{2} \\ -1 & -\frac{a}{2} & -\frac{a}{2} + 1 \end{pmatrix}.$$

La méthode précédente est très lourde. Pour calculer $\exp(A)$, il est souvent préférable de calculer les matrices A^k en fonction des premières puissances de A ou de matrices particulières ; on est alors ramené au calcul de la somme de plusieurs séries numériques.

Exemple

Soit $A = \begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}$. Calculer $\exp(A)$.

On calcule d'abord le polynôme caractéristique : $\chi_A = (X-2)^2(X-4)$.
La matrice $A - 2\mathrm{I}_3$ *étant clairement de rang* 1, *le sous-espace propre associé à la valeur propre* 2 *est de dimension* 2 *d'après le théorème du rang, et, par suite,* A *est diagonalisable.*
Son polynôme minimal est donc $\pi_A = (X-2)(X-4)$. *Pour* $n \in \mathbb{N}$, *la division euclidienne de* X^n *par* π_A *peut s'écrire :*

$$X^n = \pi_A Q + a(X-2) + b(X-4) \quad \text{avec } a,\, b \in \mathbb{R}\,.$$

✓ *Nous avons utilisé ici le fait que la famille* $(X-2, X-4)$ *est une base de* $\mathbb{R}_1[X]$. *Cette écriture permet de simplifier un peu les calculs qui suivent. Il s'agit d'un principe que nous avons déjà énoncé : en algèbre linéaire en dimension finie, le choix d'une bonne base est primordial.*

En évaluant cette relation en 2 *et en* 4, *on trouve :*

$$X^n = \pi_A Q + \frac{1}{2}\left(4^n(X-2) - 2^n(X-4)\right),$$

et en appliquant cette égalité polynomiale à la matrice A, *puisque* $\pi_A(A) = 0_3$, *on trouve :*

$$A^n = \frac{1}{2}\left(4^n(A - 2\mathrm{I}_3) - 2^n(A - 4\mathrm{I}_3)\right),$$

d'où l'on tire facilement :

$$\begin{aligned}\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!} &= \frac{1}{2}\left(\mathrm{e}^4(A - 2\mathrm{I}_3) - \mathrm{e}^2(A - 4\mathrm{I}_3)\right) \\ &= \frac{1}{2}\begin{pmatrix} \mathrm{e}^4 + \mathrm{e}^2 & -3\mathrm{e}^4 + 3\mathrm{e}^2 & 2\mathrm{e}^4 - 2\mathrm{e}^2 \\ -\mathrm{e}^4 + \mathrm{e}^2 & 3\mathrm{e}^4 - \mathrm{e}^2 & -2\mathrm{e}^4 + 2\mathrm{e}^2 \\ -\mathrm{e}^4 + \mathrm{e}^2 & 3\mathrm{e}^4 - 3\mathrm{e}^2 & -2\mathrm{e}^4 + 4\mathrm{e}^2 \end{pmatrix}.\end{aligned}$$

Exemple

Soient $a,\, b \in \mathbb{C}$. Calculer $\exp(A)$ avec $A = \begin{pmatrix} a & b & \cdots & \cdots & b \\ b & a & b & \cdots & b \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b & b & \ddots & a & b \\ b & b & b & \cdots & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$.

Nous avons déjà calculé les puissances successives de cette matrice en **[SF3.7]** *page 253. Nous pourrions donc nous servir du résultat et procéder*

comme dans l'exercice précédent, mais nous allons exposer une méthode plus rapide.
En notant J la matrice d'ordre n dont tous les coefficients sont égaux à 1, on a $A = (a-b)\mathrm{I}_n + bJ$, donc $\exp(A) = \mathrm{e}^{a-b}\exp(bJ)$.
Or $J^k = n^{k-1}J$ pour $k \geqslant 1$ donc :

$$\exp(bJ) = \mathrm{I}_n + \left(\sum_{k=1}^{+\infty} \frac{b^k n^{k-1}}{k!}\right) J = \mathrm{I}_n + \frac{1}{n}\left(\sum_{k=1}^{+\infty} \frac{(nb)^k}{k!}\right) J$$
$$= \mathrm{I}_n + \frac{1}{n}(\mathrm{e}^{nb} - 1)J\,,$$

et finalement : $\exp(A) = \mathrm{e}^{a-b}\left[\mathrm{I}_n + \frac{1}{n}(\mathrm{e}^{nb} - 1)J\right]$.

SF4.30 Étudier l'exponentielle d'une matrice

Les résultats sur la réduction des endomorphismes permettent d'obtenir un certain nombre de propriétés de l'exponentielle d'une matrice.

Exemple

Si $A \in \mathcal{M}_p(\mathbb{K})$, montrer qu'il existe $P \in \mathbb{K}_{p-1}[X]$ tel que $\exp(A) = P(A)$.

- *Si P est un polynôme quelconque de $\mathbb{K}[X]$, sa division euclidienne par χ_A s'écrit :*
$$P = Q\chi_A + R \quad \text{où } \deg(R) < \deg(\chi_A)\,,$$
donc $P(A) = Q(A)\chi_A(A) + R(A) = R(A)$ d'après le théorème de Cayley-Hamilton.
- *Ainsi, pour tout $P \in \mathbb{K}[X]$, $P(A) \in F = \mathrm{Vect}(\mathrm{I}_n, A, \ldots, A^{p-1})$. En particulier, pour tout $n \in \mathbb{N}$ les matrices $P_n(A) = \sum_{k=0}^{n} \frac{1}{k!}A^k$ appartiennent à F. F étant un sous-espace vectoriel de dimension finie, il est fermé ; donc $\exp(A) = \lim_{n\to+\infty} P_n(A) \in F$, ce qu'il fallait démontrer.*

  ✓ Le résultat ci-dessus permet de retrouver facilement la propriété vue en **[S11.10]** : si A et B commutent, alors $\exp(A+B) = \exp(A)\times\exp(B)$ (en effet, tous les polynômes en A commutent alors avec tous les polynômes en B).

Exemple

Soit $A \in \mathcal{M}_p(\mathbb{K})$. Montrer que $\det\big(\exp(A)\big) = \mathrm{e}^{\mathrm{tr}(A)}$.

Dans $\mathcal{M}_p(\mathbb{C})$ il est toujours possible de trigonaliser A : $A = PTP^{-1}$, avec P inversible et T triangulaire supérieure.
On sait que les éléments diagonaux de T sont les valeurs propres $\lambda_1, \ldots, \lambda_p$ de A (dans $\mathbb{C}$). Pour tout $k \in \mathbb{N}$, les éléments diagonaux de T^k seront donc les λ_i^k, et ceux de $\exp(T)$ seront les e^{λ_i} (toutes ces matrices sont triangulaires supérieures).
Puisque $\exp(A) = P\exp(T)P^{-1}$, on a :

$$\det\big(\exp(A)\big) = \det\big(\exp(T)\big) = \prod_{i=1}^{p} \mathrm{e}^{\lambda_i} = \exp\left(\sum_{i=1}^{p} \lambda_i\right) = \mathrm{e}^{\mathrm{tr}(A)}\,.$$

À vous de jouer !

Exercice 4.1

SF4.1 **SF4.3** **SF4.7** **SF4.11** **SF4.13**

Soit f l'endomorphisme de $\mathbb{R}^4$ de matrice $A = \begin{pmatrix} 4 & 7 & 6 & -7 \\ 2 & 3 & 1 & -3 \\ 7 & -1 & 0 & -7 \\ 4 & 7 & 5 & -7 \end{pmatrix}$ dans la base canonique.

1) Calculer le polynôme caractéristique de f.

2) Déterminer les sous-espaces propres de f. f est-il diagonalisable ?

3) On note $F = \mathrm{Ker}(f - 2\mathrm{Id})^2$ et $G = \mathrm{Ker}(f + 2\mathrm{Id})^2$. Montrer que $\mathbb{R}^4 = F \oplus G$.

4) Déterminer des bases de F et de G et écrire la matrice de f dans la nouvelle base de $\mathbb{R}^4$ associée.

Exercice 4.2

SF4.1 **SF4.5** **SF4.8**

Soit a un réel. Diagonaliser la matrice (d'ordre $n \geqslant 3$) :

$$A = \begin{pmatrix} 1 & a & a & \cdots & a & a \\ a & 1 & 0 & \cdots & 0 & a \\ a & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 & a \\ a & 0 & 0 & \cdots & 1 & a \\ a & a & a & \cdots & a & 1 \end{pmatrix}.$$

Exercice 4.3

SF4.1

Soient a et b deux entiers naturels non nuls.
Déterminer les éléments propres de l'endomorphisme u de $\mathbb{R}[X]$ qui, à tout polynôme P, associe le polynôme $u(P) = (aX + b)P + X(1 - X)P'$.

Exercice 4.4

SF4.1

Soit $E = \mathscr{C}([0\,;1],\mathbb{R})$ et $u : \begin{cases} E & \longrightarrow & E \\ f & \longmapsto & \left(x \longmapsto \displaystyle\int_0^1 \min(x,t)\, f(t)\, \mathrm{d}t\right) \end{cases}$

Déterminer les éléments propres de f.

Exercice 4.5

SF4.13

Soit u un endomorphisme d'un espace vectoriel E de dimension finie n, et λ une valeur propre de u.
Montrer que les propriétés suivantes sont équivalentes :

(i) $E = E_\lambda(u) \oplus \mathrm{Im}(u - \lambda \mathrm{Id}_E)$;

(ii) $E_\lambda(u)$ possède un supplémentaire stable par u ;

(iii) $\dim E_\lambda(u) = m_\lambda$, où m_λ est l'ordre de multiplicité de λ ;

(iv) λ est racine simple du polynôme minimal π_u de u .

Montrer que si ces conditions sont réalisées, $\mathrm{Im}(u-\lambda \mathrm{Id}_E)$ est le seul supplémentaire de $E_\lambda(u)$ stable par u.

Exercice 4.6 **SF4.3** **SF4.20**

Soit la suite $(u_n)_{n\in\mathbb{N}}$ définie par :

$$u_0 = 1 + \cos\theta\,,\ u_1 = -1\,,\ u_2 = 4 + \cos\theta\ (\theta \in]0\,;\pi[)$$

et la relation de récurrence :

$$\forall\, n \in \mathbb{N},\ u_{n+3} + 2(1-\cos\theta)u_{n+2} + (1 - 4\cos\theta)u_{n+1} + 2u_n = 0\,.$$

Calculer u_n en fonction de n.

Exercice 4.7 **SF4.2** **SF4.27**

Soit $A \in \mathcal{M}_n(\mathbb{R})$ tel que $A^3 - 3A + 4\mathrm{I}_n = 0$. Déterminer le signe de $\det A$.

Exercice 4.8 **SF4.9**

Soient A et B deux matrices symétriques réelles telles que $A^3 = B^3$.
Prouver que $A = B$.

Exercice 4.9 **SF4.2** **SF4.3** **SF4.10**

n désigne un entier naturel non nul. Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $B = \begin{bmatrix} 0 & \mathrm{I}_n \\ A & 0 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$.

1) Calculer le polynôme caractéristique de B en fonction de celui de A.

2) Montrer que : B diagonalisable $\iff$ A diagonalisable et inversible.

3) Lorsque $A \in \mathcal{M}_n(\mathbb{R})$, trouver une condition nécessaire et suffisante pour que B soit diagonalisable dans $\mathcal{M}_{2n}(\mathbb{R})$.

Exercice 4.10 **SF4.2** **SF4.10**

Soit $A \in \mathcal{M}_n(\mathbb{K})$ fixée. On note f_A et g_A les applications définies sur $\mathcal{M}_n(\mathbb{K})$ par

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}),\ f_A(M) = AM \quad \text{et} \quad g_A(M) = MA\,.$$

1) Montrer que f_A est inversible si et seulement si A l'est.

2) a) Montrer que f_A et A ont même spectre.

b) Montrer que f_A est diagonalisable si et seulement si A l'est.

3) Montrer que : A diagonalisable $\Longrightarrow$ g_A diagonalisable.

4) Soient A et B deux matrices diagonalisables de $\mathcal{M}_n(\mathbb{K})$. Montrer que l'application u, qui à toute matrice M associe la matrice $AM+MB$, est diagonalisable.

Exercice 4.11

SF4.1 **SF4.10**

Soient A, $B \in \mathcal{M}_n(\mathbb{C})$ telles que $AB - BA = A$.

1) Montrer que : $\forall\, k \in \mathbb{N}$, $A^k B - BA^k = kA^k$.

2) En considérant l'endomorphisme $\varphi\colon M \mapsto MB - BM$, montrer que A est nilpotente.

Exercice 4.12

SF4.28

Montrer que, si f est un endomorphisme d'un $\mathbb{R}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, il existe une droite ou un plan stable par f.

Exercice 4.13

SF4.22 **SF4.28**

Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et u un endomorphisme de E ayant n valeurs propres distinctes.

Montrer que $\mathscr{C}(u) = \mathbb{K}[u]$ ($\mathscr{C}(u)$ désigne le commutant de u).

Exercice 4.14

SF4.1 **SF4.2** **SF4.3**

Soient A et B deux matrices carrées d'ordre n à coefficients dans $\mathbb{C}$.

1) On suppose ici que A et B n'ont pas de valeur propre commune.

a) Soit χ_A le polynôme caractéristique de A. Montrer que $\chi_A(B)$ est inversible.

b) Montrer que, pour tout $X \in \mathcal{M}_n(\mathbb{C})$, on a : $AX = XB \Longleftrightarrow X = 0$.

c) Montrer que, pour tout $M \in \mathcal{M}_n(\mathbb{C})$, il existe une et une seule matrice $X \in \mathcal{M}_n(\mathbb{C})$ telle que $AX - XB = M$.

2) On suppose ici que A et B ont au moins une valeur propre commune λ. Montrer que λ est valeur propre de ${}^tB$. En déduire une matrice $X \in \mathcal{M}_n(\mathbb{C})$ non nulle telle que $AX = XB$.

3) On suppose ici qu'il existe $X \in \mathcal{M}_n(\mathbb{C})$, non nulle, telle que $AX = XB$. Soit $r = \operatorname{rg} X$. Montrer que A et B ont au moins r valeurs propres communes (comptées avec leur ordre de multiplicité).

■ Comment montrer qu'une application est un produit scalaire ?

SF5.1 Revenir à la définition

Il s'agit de la méthode la plus courante : pour montrer que $\varphi\colon E \times E \to \mathbb{R}$ est un produit scalaire, on vérifie point par point qu'il s'agit d'une forme bilinéaire symétrique définie positive.
Généralement, la bilinéarité et la symétrie sont évidentes ; il faut par contre montrer avec soin le caractère positif et surtout défini. De plus, il ne faut pas oublier de vérifier que φ est bien une application sur $E \times E$, c'est-à-dire que l'expression donnée a bien un sens.

Exemple

À quelle condition l'application :

$$\varphi\colon \big((x,y),(x',y')\big) = axx' + bxy' + cx'y + dyy'$$

est-elle un produit scalaire sur $\mathbb{R}^2$?

- *φ est bien définie sur $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.*
- *Notons $e_1 = (1,0)$ et $e_2 = (0,1)$ les vecteurs de la base canonique de $\mathbb{R}^2$. Si φ est symétrique, alors $\varphi(e_1,e_2) = \varphi(e_2,e_1)$, ce qui implique $b = c$, et réciproquement, si $b = c$ il est clair que pour tout $(x,y,x',y') \in \mathbb{R}^4$ on a $\varphi\big((x,y),(x',y')\big) = \varphi\big((x',y'),(x,y)\big)$.*
  En conclusion : φ symétrique $\iff b = c$. On supposera cette condition réalisée désormais.
- *Pour (x',y') fixé dans $\mathbb{R}^2$, l'application $(x,y) \mapsto \varphi\big((x,y),(x',y')\big)$ est linéaire puisqu'elle est de la forme $(x,y) \mapsto Ax + By$.*
  Ainsi φ est linéaire à gauche et symétrique, donc bilinéaire.
- *Pour tout $(x,y) \in \mathbb{R}^2$ on calcule :*

$$\varphi\big((x,y),(x,y)\big) = ax^2 + 2bxy + dy^2\,.$$

  - *Si $y = 0$, cette expression est strictement positive pour tout $x \neq 0$ si et seulement si $a > 0$, ce que l'on supposera par la suite.*
  - *Pour $y \neq 0$,*

$$\varphi\big((x,y),(x,y)\big) = y^2\left(a\left(\frac{x}{y}\right)^2 + 2b\frac{x}{y} + d\right) = y^2(at^2 + 2bt + d),$$

    en ayant posé $t = \dfrac{x}{y}$, et puisque $a > 0$, cette expression est strictement positive pour tout t si et seulement si $b^2 - ad < 0$ (le trinôme n'a pas de racine).

En conclusion, φ est un produit scalaire si et seulement si :

$$b = c\ ,\ a > 0 \text{ et } b^2 - ad < 0\,.$$

Exemple

Montrer que l'application : φ: $(P,Q) \mapsto \sum_{n=0}^{+\infty} P(n)Q(n)\mathrm{e}^{-n}$ définit un produit scalaire sur $\mathbb{R}[X]$.

- *Il faut commencer ici par vérifier que cette définition a bien un sens, c'est-à-dire que la série écrite converge.*
  Or, si P et Q sont deux polynômes, d'après les croissances comparées des fonctions usuelles : $\lim_{n\to+\infty} n^2 P(n)Q(n)\mathrm{e}^{-n} = 0$, donc $P(n)Q(n)\mathrm{e}^{-n} \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$; la série à termes positifs $\sum \frac{1}{n^2}$ étant convergente, il résulte des théorèmes de comparaison sur les séries à termes positifs que la série de terme général $P(n)Q(n)\mathrm{e}^{-n}$ est absolument convergente donc convergente.
  φ est donc bien une application de $\mathbb{R}[X]^2$ dans $\mathbb{R}$.
- *L'égalité $\varphi(P,Q) = \varphi(Q,P)$ est immédiate : φ est symétrique.*
- *La linéarité de l'application $P \mapsto \varphi(P,Q)$ ne pose pas de problèmes, compte tenu des propriétés des opérations sur les séries convergentes. Il s'ensuit que φ est linéaire à gauche, puis bilinéaire.*
- *Pour tout $P \in \mathbb{R}[X]$, $\varphi(P,P) = \sum_{n=0}^{+\infty} P(n)^2\mathrm{e}^{-n}$ est positif, de façon claire. Donc φ est positive.*
- *Enfin, si $\varphi(P,P) = 0$ alors $P(n) = 0$ pour tout $n \in \mathbb{N}$; P est alors un polynôme qui possède une infinité de racines, donc $P = 0$: cela prouve que φ est définie, et achève la démonstration.*

SF5.2 Utiliser les matrices

Cette méthode, qui ne peut être utilisée que si l'espace vectoriel E est de dimension finie, est particulièrement performante et évite de nombreux calculs dans beaucoup de cas. Nous en exposons le principe ci-dessous.

> ✓ La notion de matrice d'une forme bilinéaire n'étant plus au programme, il faut savoir refaire ces calculs. Des précisions sur les notions utilisées ici seront données dans l'exercice **6.8**.

Soit E un $\mathbb{R}$-espace vectoriel de dimension n, rapporté à une base $\mathscr{B} = (e_1,\dots,e_n)$. Si φ est une application bilinéaire de $E \times E$ dans $\mathbb{R}$, on a, si $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$ sont deux vecteurs de E, en utilisant la bilinéarité de φ :

$$\varphi(x,y) = \varphi\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{1\leqslant i,j\leqslant n} x_i y_j \varphi(e_i,e_j)\,,$$

ce qui peut aussi s'écrire, en notant $a_{ij} = \varphi(e_i,e_j)$ pour tout $(i,j) \in [\![1\,;n]\!]^2$:

$$\varphi(x,y) = \sum_{1\leqslant i,j\leqslant n} a_{ij} x_i y_j \qquad (*)$$

Réciproquement, il est facile de vérifier que si φ est l'application de $E \times E$ dans $\mathbb{R}$ donnée par la relation $(*)$, alors φ est une forme bilinéaire.

La matrice $A = (a_{ij})$ s'appelle la matrice de φ dans la base $\mathscr{B}$. Si l'on note $X = {}^{\mathrm{t}}(x_1 \quad \dots \quad x_n)$ et $Y = {}^{\mathrm{t}}(y_1 \quad \dots \quad y_n)$ les matrices colonnes des coordonnées de x et y dans la base $\mathscr{B}$, la relation $(*)$ peut aussi s'écrire sous forme matricielle :

$$\varphi(x,y) = {}^{\mathrm{t}}XAY\,.$$

Les propriétés de φ découlent alors de celles de A ; en particulier :

- φ est symétrique si et seulement si A est symétrique ;
- φ est positive si et seulement si pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$ on a ${}^{\mathrm{t}}XAX \geqslant 0$, ce qui peut se démontrer en diagonalisant la matrice A qui est symétrique réelle.

Exemple

À quelle condition portant sur les réels a et b l'application φ qui à deux vecteurs $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ associe :

$$\varphi(x,y) = a\sum_{i=1}^{n} x_i y_i + b\sum_{i\neq j} x_i y_j\,,$$

définit-elle un produit scalaire sur $\mathbb{R}^n$?

• *Dans le cas $b = 0$, $\varphi(x,y) = a\langle x|y\rangle$ où $\langle x|y\rangle$ est le produit scalaire canonique sur $\mathbb{R}^n$ donc φ sera un produit scalaire si et seulement si $a > 0$.*

• *Pour la suite on supposera donc $b \neq 0$.*
On a pour $\varphi(x,y)$ une expression de la forme $(*)$ donc φ est bien une forme bilinéaire sur $\mathbb{R}^n$, dont la matrice associée est :

$$A = \begin{pmatrix} a & b & \dots & \dots & b \\ b & a & b & \dots & b \\ b & \ddots & \ddots & \ddots & b \\ b & b & \ddots & a & b \\ b & b & b & \dots & a \end{pmatrix}.$$

Cette matrice étant symétrique, φ est symétrique puisque, avec les notations habituelles :

$$\varphi(y,x) = {}^{\mathrm{t}}YAX = {}^{\mathrm{t}}({}^{\mathrm{t}}YAX) = {}^{\mathrm{t}}X{}^{\mathrm{t}}AY = {}^{\mathrm{t}}XAY = \varphi(x,y)\,.$$

✓ *On a utilisé dans le calcul ci-dessus une astuce bien utile en algèbre bilinéaire : puisque ${}^{\mathrm{t}}YAX$ est un réel, il est égal à sa transposée.*

La matrice A étant symétrique réelle est ortho-diagonalisable ; cherchons ses valeurs propres (on n'a pas besoin de déterminer les sous-espaces propres, comme on le verra plus loin). Pour cela on remarque que la matrice $A + (b-a)I_n$ est de rang 1 (car $b \neq 0$) donc $\dim \operatorname{Ker}(A + (b-a)I_n) = n-1$ ce qui signifie que $a - b$ est valeur propre de A d'ordre $n-1$ (car A étant diagonalisable, la dimension de chaque sous-espace propre est égale à l'ordre de multiplicité de la valeur propre correspondante).
Puisque $\operatorname{tr} A = na$, la dernière valeur propre de A est $a + (n-1)b$.

Il existe donc $P \in \mathcal{O}_n(\mathbb{R})$ (soit $P^{-1} = {}^tP$) telle que $A = PDP^{-1}$, avec :

$$D = \mathrm{diag}(\underbrace{a-b,\ldots,a-b}_{n-1 \text{ fois}},a+(n-1)b)\,.$$

Soit alors $x \in \mathbb{R}^n$, $X \in \mathcal{M}_{n,1}(\mathbb{R})$ la matrice colonne de ses coordonnées dans la base canonique, et $Y = P^{-1}X$ (Y représente en fait les coordonnées de x dans la base de vecteurs propres définie par la matrice de passage P). On a donc :

$$\varphi(x,x) = {}^tXAX = {}^t(PY)A(PY) = {}^tY{}^tPAPY = {}^tYDY\,,$$

donc si l'on note $Y = {}^t\begin{pmatrix} y_1 & \ldots & y_n \end{pmatrix}$:

$$\varphi(x,x) = (a-b)y_1^2 + \cdots + (a-b)y_{n-1}^2 + (a+(n-1)b)y_n^2\,.$$

Sous cette forme il est clair que φ est définie positive si et seulement si $a-b>0$ et $a+(n-1)b>0$.

• *En conclusion (on peut réintégrer le cas $b=0$), φ est un produit scalaire si et seulement si : $a-b>0$ et $a+(n-1)b>0$.*

■ Comment utiliser un produit scalaire ?

SF5.3 Inégalité de Cauchy-Schwarz

Il faut y penser dès que l'on a une inégalité à démontrer dans un espace préhilbertien, mais avant de l'utiliser, il ne faut pas oublier de préciser clairement :

– dans quel espace on se place ;

– quel est le produit scalaire considéré ;

– quels sont les deux vecteurs auxquels on applique l'inégalité.

Exemple

Soient x, y, z des réels tels que $2x^2+y^2+3z^2 \leqslant 1$.

Montrer que $(x+y+z)^2 \leqslant \dfrac{11}{6}$.

L'idée est ici de considérer le produit scalaire défini sur $\mathbb{R}^3$ par :

$$\forall\, u=(x,y,z) \in \mathbb{R}^3,\ \forall\, v=(x',y',z') \in \mathbb{R}^3,\ \varphi(u,v) = 2xx'+yy'+3zz'\,.$$

La vérification du fait qu'il s'agit bien d'un produit scalaire sur $\mathbb{R}^3$ est aisée, par l'une ou l'autre des méthodes exposées précédemment.
L'inégalité de Cauchy-Schwarz pour φ s'écrit alors, avec les mêmes notations que ci-dessus :

$$\varphi(u,v)^2 \leqslant (2x^2+y^2+3z^2)(2x'^2+y'^2+3z'^2)\,.$$

Donc si u est tel que $2x^2+y^2+3z^2 \leqslant 1$ et si on applique l'inégalité avec $v = (x',y',z') = \left(\dfrac{1}{2},1,\dfrac{1}{3}\right)$, de sorte que $\varphi(u,v) = x+y+z$, on obtient le résultat demandé.

Exemple

On munit $\mathcal{M}_n(\mathbb{R})$ du produit scalaire canonique défini par :

$$\forall\,(A,B)\in\mathcal{M}_n(\mathbb{R})^2,\ \langle A\,|\,B\rangle=\operatorname{tr}({}^tAB)\,.$$

1) Montrer que :

$$\forall\,A\in\mathcal{M}_n(\mathbb{R}),\ \operatorname{tr}A\leqslant\sqrt{n\operatorname{tr}({}^tAA)}$$

et préciser les cas d'égalité.

2) Montrer que :

$$\forall\,(A,B)\in\mathcal{M}_n(\mathbb{R})^2,\ \|AB\|\leqslant\|A\|\,\|B\|\,.$$

1) • *On applique l'inégalité de Cauchy-Schwarz aux matrices I_n et A :*

$$|\langle I_n\,|\,A\rangle|\leqslant\|I_n\|\,\|A\|\,.$$

Puisque $\|I_n\|=\sqrt{\operatorname{tr}(I_n)}=\sqrt{n}$ on obtient : $|\operatorname{tr}A|\leqslant\sqrt{n\operatorname{tr}({}^tAA)}$ ce qui implique l'inégalité demandée.

• *Il y a égalité si et seulement si $\operatorname{tr}A=|\operatorname{tr}A|$ et si il y a égalité dans l'inégalité de Cauchy-Schwarz. Cette dernière condition équivaut à l'existence d'un réel λ tel que $A=\lambda I_n$. La 1re condition donne en plus $\lambda\geqslant 0$.*

2) *En notant $A=(a_{ij})_{1\leqslant i,j\leqslant n}$, $B=(b_{ij})_{1\leqslant i,j\leqslant n}$ et $C=AB=(c_{ij})_{1\leqslant i,j\leqslant n}$, la formule du produit matriciel donne :*

$$\forall\,(i,j)\in[\![1\,;n]\!],\ c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}\,.$$

L'inégalité de Cauchy-Schwarz appliqué au produit scalaire canonique dans $\mathbb{R}^n$ pour les vecteurs $u=(a_{i1},\dots,a_{in})$ et $v=(b_{1j},\dots,b_{nj})$ donne :

$$c_{ij}^2=\langle u\,|\,v\rangle^2\leqslant\|u\|^2\,\|v\|^2=\left(\sum_{k=1}^{n}a_{ik}^2\right)\left(\sum_{\ell=1}^{n}b_{\ell j}^2\right)\,,$$

d'où :

$$\|C\|^2=\sum_{1\leqslant i,j\leqslant n}c_{ij}^2\leqslant\sum_{1\leqslant i,j\leqslant n}\left(\sum_{k=1}^{n}a_{ik}^2\right)\left(\sum_{\ell=1}^{n}b_{\ell j}^2\right)$$

$$\leqslant\underbrace{\left(\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}^2\right)}_{=\|A\|^2}\underbrace{\left(\sum_{j=1}^{n}\sum_{\ell=1}^{n}b_{\ell j}^2\right)}_{=\|B\|^2}$$

puis l'inégalité demandée.

✓ *Notez que dans tous les calculs précédents, nous avons considéré les normes au carré : les écritures sont simplifiées puisqu'il n'y a pas de radicaux.*

SF5.4 Utiliser les identités de polarisation

La norme est naturellement définie à partir du produit scalaire par la relation $\|x\|^2=\langle x\,|\,x\rangle$, mais il ne faut pas oublier que l'on peut aussi retrouver le produit scalaire à partir de la norme grâce aux identités de polarisation **[S5.1]**.

Exemple

Soit E un espace préhilbertien et u une application de E dans E telle que :

$$u(0)=0 \quad \text{et} \quad \forall\,(x,y)\in E^2,\ \|u(x)-u(y)\|=\|x-y\|\ .$$

Montrer que u est linéaire.

En faisant $y=0$ dans la relation, on a $\|u(x)\|=\|x\|$ pour tout x.
Pour tous x, $y\in E$, on a alors, à l'aide des identités de polarisation :

$$\begin{aligned}\langle u(x)\,|\,u(y)\rangle &= -\frac{1}{2}\left(\|u(x)-u(y)\|^2-\|u(x)\|^2-\|u(y)\|^2\right)\\ &= -\frac{1}{2}\left(\|x-y\|^2-\|x\|^2-\|y\|^2\right)=\langle x\,|\,y\rangle\ .\end{aligned}$$

Ainsi u conserve le produit scalaire.
Soient maintenant $\lambda\in\mathbb{R}$ et x, $y\in E$. On calcule :

$$\begin{aligned}&\|u(\lambda x+y)-\lambda u(x)-u(y)\|^2\\ &\quad=\|u(\lambda x+y)\|^2+\lambda^2\|u(x)\|^2+\|u(y)\|^2+2\lambda\langle u(x)\,|\,u(y)\rangle\\ &\qquad\qquad-2\lambda\langle u(\lambda x+y)\,|\,u(x)\rangle-2\langle u(\lambda x+y)\,|\,u(y)\rangle\\ &\quad=\|\lambda x+y\|^2+\lambda^2\|x\|^2+\|y\|^2+2\lambda\langle x\,|\,y\rangle\\ &\qquad\qquad-2\lambda\langle\lambda x+y\,|\,x\rangle-2\langle\lambda x+y\,|\,y\rangle\\ &\quad=\|(\lambda x+y)-\lambda x-y\|^2=0\,.\end{aligned}$$

Ainsi $\|u(\lambda x+y)-\lambda u(x)-u(y)\|^2=0$ d'où $u(\lambda x+y)-\lambda u(x)-u(y)=0$ c'est-à-dire u linéaire.

La même idée peut être appliquée dans l'exemple suivant, même s'il ne s'agit pas d'un produit scalaire : à partir de valeurs de $\varphi(x,x)$ (la norme dans le cas d'un produit scalaire), on peut obtenir des renseignements sur $\varphi(x,y)$ pour tous x, y.

Exemple

Soit E un $\mathbb{R}$-espace vectoriel et $\varphi\colon E\times E\to\mathbb{R}$ une forme bilinéaire sur E. On suppose que :

$$\forall\,x\in E,\ \varphi(x,x)=0\,.$$

Montrer alors que : $\forall\,(x,y)\in E^2$, $\varphi(y,x)=-\varphi(x,y)$ (φ est dite antisymétrique).

Il suffit d'appliquer la relation donnée à $x+y$:

$$\varphi(x+y,x+y)=0\,.$$

En utilisant la bilinéarité de φ on obtient :

$$\varphi(x,x)+\varphi(x,y)+\varphi(y,x)+\varphi(y,y)=0\,,$$

et puisque $\varphi(x,x)=\varphi(y,y)=0$, on obtient le résultat demandé.

SF5.5 Utiliser la norme

Outre l'inégalité de Cauchy-Schwarz, il faut aussi penser à utiliser l'inégalité triangulaire, la relation de Pythagore et l'identité du parallélogramme.

Exemple

Dans un espace préhilbertien E, montrer que :

$$\forall\,(x,y)\in E^2,\ \left|\|x\|^2-\|y\|^2\right|\leqslant\|x+y\|\,.\,\|x-y\|\leqslant\|x\|^2+\|y\|^2\ .$$

On calcule le carré du terme central :

$$\begin{aligned}\|x+y\|^2\,.\,\|x-y\|^2&=\left(\|x\|^2+\|y\|^2+2\,\langle x\,|\,y\rangle\right)\left(\|x\|^2+\|y\|^2-2\,\langle x\,|\,y\rangle\right)\\&=\left(\|x\|^2+\|y\|^2\right)^2-4\,\langle x\,|\,y\rangle^2\\&=\|x\|^4+\|y\|^4+2\,\|x\|^2\,.\,\|y\|^2-4\,\langle x\,|\,y\rangle^2\\&\leqslant\|x\|^4+\|y\|^4+2\,\|x\|^2\,.\,\|y\|^2=\left(\|x\|^2+\|y\|^2\right)^2\end{aligned}$$

ce qui fournit l'inégalité de droite.
Le calcul montre aussi qu'il y a égalité dans l'inégalité de droite si et seulement si $\langle x\,|\,y\rangle=0$ c'est-à-dire si x et y sont orthogonaux.
Notons que cette inégalité pouvait aussi être obtenue à l'aide de l'identité du parallélogramme puisque :

$$\begin{aligned}&\|x\|^2+\|y\|^2-\|x+y\|\,.\,\|x-y\|\\&\qquad=\frac{1}{2}\left(\|x+y\|^2+\|x-y\|^2\right)-\|x+y\|\,.\,\|x-y\|\\&\qquad=\frac{1}{2}\left(\|x+y\|-\|x-y\|\right)^2\geqslant 0\,.\end{aligned}$$

Pour l'inégalité de gauche, reprenons le premier calcul :

$$\begin{aligned}\|x+y\|^2\,.\,\|x-y\|^2&=\|x\|^4+\|y\|^4+2\,\|x\|^2\,.\,\|y\|^2-4\,\langle x\,|\,y\rangle^2\\&=\|x\|^4+\|y\|^4-2\,\|x\|^2\,.\,\|y\|^2+4\left(\|x\|^2\,.\,\|y\|^2-\langle x\,|\,y\rangle^2\right)\\&=\left(\|x\|^2-\|y\|^2\right)^2+4\left(\|x\|^2\,.\,\|y\|^2-\langle x\,|\,y\rangle^2\right)\ .\end{aligned}$$

Or d'après l'inégalité de Cauchy-Schwarz, $\langle x\,|\,y\rangle^2\leqslant\|x\|^2\,.\,\|y\|^2$, donc

$$\|x+y\|^2\,.\,\|x-y\|^2\geqslant\left(\|x\|^2-\|y\|^2\right)^2\ ,$$

ce qui donne l'inégalité cherchée. De plus, il y a égalité si et seulement si le système $\{x,y\}$ est lié (cas d'égalité dans l'inégalité de Cauchy-Schwarz).

SF5.6 Montrer que deux vecteurs sont égaux

La méthode que nous exposons ici, et qui peut sembler de prime abord un peu alambiquée, est une des méthodes fondamentales de l'algèbre bilinéaire.

Pour montrer que deux vecteurs x et y sont égaux, on est ramené à démontrer que le vecteur $u=x-y$ est nul. Pour cela on peut alors :

- montrer que $\|u\|=0$, ou, mieux, montrer que $\|u\|^2=0$; c'est cette méthode que nous avons utilisée dans le premier exemple de **[SF5.4]** ;
- montrer que u est orthogonal à tout l'espace (puisque $E^\perp=\{0\}$, c'est-à-dire montrer que pour tout $z\in E$, $\langle u\,|\,z\rangle=0$.

Exemple

Soit E préhilbertien réel, et f, g deux applications de E dans E telles que :

$$\forall\,(x,y)\in E^2,\ \langle f(x)\,|\,y\rangle = \langle x\,|\,g(y)\rangle\ .$$

Montrer que f et g sont linéaires.

Le problème posé étant symétrique en f et g, il suffit de montrer que f est linéaire, c'est-à-dire que pour tout $\lambda\in\mathbb{R}$ et tous x, $x'\in E$, on a :

$$f(\lambda x+x') = \lambda f(x)+f(x')\,,$$

ce qui équivaut, d'après la remarque précédente, à montrer que :

$$\forall\, z\in E,\ \big\langle f(\lambda x+x')-\lambda f(x)-f(x')\,\big|\,z\big\rangle = 0\,.$$

Il suffit alors de calculer, en utilisant la propriété donnée et la bilinéarité du produit scalaire :

$$\begin{aligned}\big\langle f(\lambda x+x')-\lambda f(x)-f(x')\,\big|\,z\big\rangle &= \langle f(\lambda x+x')\,|\,z\rangle - \lambda\,\langle f(x)\,|\,z\rangle - \langle f(x')\,|\,z\rangle\\ &= \langle \lambda x+x'\,|\,g(z)\rangle - \lambda\,\langle x\,|\,g(z)\rangle - \langle x'\,|\,g(z)\rangle\\ &= 0\,,\end{aligned}$$

et cela achève la démonstration.

Exemple

Soit $A\in\mathcal{M}_n(\mathbb{R})$. Montrer que $\operatorname{Ker} A = \operatorname{Ker}{}^tAA$, et en déduire $\operatorname{Im}{}^tAA = \operatorname{Im}{}^tA$.

- *Pour tout $X\in\mathcal{M}_{n,1}(\mathbb{R})$, $AX=0\Longrightarrow {}^tAAX=0$ donc* $\operatorname{Ker} A\subset\operatorname{Ker}{}^tAA$.
- *Réciproquement, soit $X\in\operatorname{Ker}{}^tAA$.*
  Alors ${}^tAAX=0$ d'où ${}^tX{}^tAAX=0$ soit ${}^t(AX)AX=0$ ou encore $\|AX\|^2=0$ (il s'agit ici de la norme dans $\mathcal{M}_{n,1}(\mathbb{R})$ identifié à $\mathbb{R}^n$ muni de son produit scalaire canonique).
  Il en résulte $AX=0$ donc $X\in\operatorname{Ker} A$ puis l'inclusion $\operatorname{Ker}{}^tAA\subset\operatorname{Ker} A$.
- *On a donc démontré :* $\operatorname{Ker} A=\operatorname{Ker}{}^tAA$.
  On déduit alors du théorème du rang que $\operatorname{rg} A=\operatorname{rg}{}^tAA$ *donc aussi* $\operatorname{rg}{}^tA=\operatorname{rg}{}^tAA$. *Puisque l'on a l'inclusion facile* $\operatorname{Im}{}^tAA\subset\operatorname{Im}{}^tA$, *l'égalité des dimensions implique l'égalité des sous-espaces vectoriels.*

  ✓ Cet exemple montre que, dans certains exercices d'algèbre linéaire en dimension finie, il ne faut pas hésiter à introduire une structure euclidienne même si elle n'apparaît pas dans l'énoncé.

■ Comment trouver l'orthogonal d'un sous-espace vectoriel ?

SF5.7 Revenir à la définition

On utilise ici simplement la définition de l'orthogonal :

$$A^\perp = \{x\in E\mid \forall\, a\in A,\ \langle a|x\rangle = 0\}\ .$$

Il s'agit d'ailleurs de la seule méthode lorsque E est de dimension infinie.

Exemple

Soit $E = \mathscr{C}([0\,;1],\mathbb{R})$ muni du produit scalaire classique : $\langle f\,|\,g\rangle = \displaystyle\int_0^1 f(t)g(t)\,\mathrm{d}t$.

Soit $H = \{f \in E \mid f(0) = 0\}$. Montrer que H est un sous-espace vectoriel de E et en déterminer l'orthogonal.

- *H est un sous-espace vectoriel de E, car c'est le noyau de la forme linéaire $f \mapsto f(0)$ (c'est un hyperplan).*
- *Soit $g \in H^\perp$; alors, pour toute $f \in H$, $\displaystyle\int_0^1 f(t)g(t)\,\mathrm{d}t = 0$.*
  Mais l'application $f\colon t \mapsto tg(t)$ appartient à H, donc on a $\displaystyle\int_0^1 tg(t)^2\,\mathrm{d}t = 0$, et, puisqu'il s'agit d'une fonction continue positive, on en déduit : $\forall t \in [0\,;1]$, $tg(t)^2 = 0$ donc $g(t) = 0$ pour $t \in]0\,;1]$ et enfin $g(t) = 0$ pour tout $t \in [0\,;1]$ puisque g est continue.
  Ainsi, $H^\perp = \{0\}$.
  ✓ On a ainsi exhibé un exemple de sous-espace vectoriel H d'un espace préhilbertien E tel que $\left(H^\perp\right)^\perp \neq H$.

SF5.8 Prouver une inclusion et une somme

Pour montrer que F et G sont supplémentaires orthogonaux, on peut se contenter de vérifier :

$$\forall\,(x,y) \in F \times G,\ \langle x\,|\,y\rangle = 0\,,$$

ce qui montre l'inclusion $G \subset F^\perp$, et implique que la somme $F + G$ est directe, puis vérifier $E = F + G$.

Exemple

Soit $E = \mathscr{C}^2([0\,;1],\mathbb{R})$. On définit sur E^2 l'application :

$$\varphi\colon (f,g) \mapsto \int_0^1 \big(f(t)g(t) + f'(t)g'(t)\big)\,\mathrm{d}t\,.$$

1) Montrer que φ est un produit scalaire sur E.

2) Soient $F = \{f \in E \mid f(0) = f(1) = 0\}$ et $G = \{g \in E \mid g'' = g\}$. Montrer que F et G sont des supplémentaires orthogonaux pour le produit scalaire φ.

1) *Montrer que φ est un produit scalaire sur E ne devrait pas poser de problème.*

2) • *Soient $f \in F$ et $g \in G$. Alors, puisque $g = g''$:*

$$\varphi(f,g) = \int_0^1 \big(f(t)g''(t) + f'(t)g'(t)\big)\,\mathrm{d}t = \big[f(t)g'(t)\big]_0^1 = 0$$

puisque $f(0) = f(1) = 0$. F et G sont donc orthogonaux.

• *Pour montrer que $E = F + G$ on cherche à écrire toute application h de classe $\mathscr{C}^2$ sur $[0\,;1]$ sous la forme $h = f + g$ avec $f \in F$ et $g \in G$.*
Les relations $f(0) = f(1) = 0$ imposent $g(0) = h(0)$ et $g(1) = h(1)$.
Or les éléments de G sont les applications de la forme $\lambda \operatorname{ch} x + \mu \operatorname{sh} x$ donc on doit avoir $\lambda = h(0)$ puis $\mu = \dfrac{h(1) - h(0)\operatorname{ch} 1}{\operatorname{sh} 1}$.

Réciproquement, il est clair que les applications :

$$g : x \mapsto h(0)\operatorname{ch} x + \frac{h(1) - h(0)\operatorname{ch} 1}{\operatorname{sh} 1}\operatorname{sh} x \quad \textit{et} \quad f = h - g$$

vérifient bien : $f \in F$, $g \in G$ et $h = f + g$.

On a ainsi prouvé : $E = F \overset{\perp}{\oplus} G$.

SF5.9 Prouver une inclusion et utiliser les dimensions

En dimension finie, il s'agit certainement de la méthode la plus rapide : pour montrer que $G = F^\perp$, on peut se contenter de montrer une inclusion, donc vérifier

$$\forall\,(x,y) \in F \times G,\ \langle x \,|\, y\rangle = 0$$

ce qui prouve que $G \subset F^\perp$, puis vérifier $\dim G = \dim E - \dim F$.
Cela prouvera $\dim G = \dim F^\perp$ donc finalement $G = F^\perp$.

Exemple

Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, montrer que les sous-espaces vectoriels $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ formés respectivement des matrices symétriques et antisymétriques sont supplémentaires orthogonaux.

- *On sait déjà que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{R})$ et qu'ils sont supplémentaires (voir* **[S3.5]***).*
- *Soit $S \in \mathcal{S}_n(\mathbb{R})$ et $A \in \mathcal{A}_n(\mathbb{R})$.*
  Alors $\langle A \,|\, S\rangle = \operatorname{tr}({}^tAS) = \operatorname{tr}(-AS) = -\operatorname{tr}(AS)$ et, d'autre part, $\langle S \,|\, A\rangle = \operatorname{tr}({}^tSA) = \operatorname{tr}(SA)$.
  Puisque $\operatorname{tr}(AS) = \operatorname{tr}(SA)$ et $\langle A \,|\, S\rangle = \langle S \,|\, A\rangle$, on en déduit $\langle A \,|\, S\rangle = -\langle A \,|\, S\rangle$ d'où $\langle A \,|\, S\rangle = 0$ puis le résultat cherché.

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Montrer que $(\operatorname{Ker} A)^\perp = \operatorname{Im} {}^tA$.

- *Soit $Y \in \operatorname{Im} {}^tA$. Il existe $X \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $Y = {}^tAX$. Donc pour tout $Z \in \operatorname{Ker} A$ on a, puisque $AZ = 0$:*

$$\langle Y \,|\, Z\rangle = {}^tYZ = {}^t({}^tAX)Z = {}^tXZ = 0\,,$$

  ce qui montre que $Y \in (\operatorname{Ker} A)^\perp$ et prouve l'inclusion $\operatorname{Im} {}^tA \subset (\operatorname{Ker} A)^\perp$.
- *On sait que $\operatorname{rg} {}^tA = \operatorname{rg} A$ donc, en utilisant le théorème du rang,*

$$\dim(\operatorname{Im} {}^tA) = \dim(\operatorname{Im} A) = n - \dim(\operatorname{Ker} A) = \dim(\operatorname{Ker} A)^\perp\,,$$

  d'où l'égalité cherchée.

SF5.10 Utiliser des bases

D'après la propriété $A^\perp = (\operatorname{Vect}(A))^\perp$, si $(e_1, \ldots, e_p)$ est une base de F, un vecteur x appartient à $F^\perp$ si et seulement si il est orthogonal à tous les e_i.

Exemple

Dans $\mathbb{R}^4$, déterminer l'orthogonal du sous-espace vectoriel

$$F = \left\{(x,y,z,t) \in \mathbb{R}^4 \;\middle|\; x + y + z + t = x - y + z - t = 0\right\}\,.$$

- *F est l'intersection de deux hyperplans distincts de $\mathbb{R}^4$, c'est donc un sous-espace vectoriel de $\mathbb{R}^4$, de dimension 2.*
  En soustrayant et en additionnant les deux équations, on voit que :
  $F = \{(x,y,z,t) \in \mathbb{R}^4 \mid z = -x \text{ et } t = -y\} = \{(x,y,-x,-y) \mid (x,y) \in \mathbb{R}^2\}$
  donc les vecteurs $e_1 = (1,0,-1,0)$ et $e_2 = (0,1,0,-1)$ forment une base de F.

  ✓ *Les raisonnements et calculs ci-dessus sont classiques (cf.* **[SF2.5]** *et* **[SF2.21]***) et doivent être parfaitement maîtrisés.*
- *Donc $F^\perp = \{u \in \mathbb{R}^4 \mid \langle u|e_1\rangle = \langle u|e_2\rangle = 0\}$. Si $u = (x,y,z,t)$, on a donc en calculant les produits scalaires :*
  $$u \in F^\perp \iff x - z = y - t = 0$$
  ce qui donne directement un système d'équations de $F^\perp$.

Exemple

Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, trouver l'orthogonal du sous-espace vectoriel $\mathcal{D}_n(\mathbb{R})$ formé des matrices diagonales.

Une base de $\mathcal{D}_n(\mathbb{R})$ est formée des $E_{i,i}$ pour $1 \leqslant i \leqslant n$ (où $(E_{i,j})$ est la base canonique de $\mathcal{M}_n(\mathbb{R})$).

Donc $\mathcal{D}_n(\mathbb{R})^\perp = \bigcap_{i=1}^{n} \{E_{i,i}\}^\perp$. Or si $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, $\langle A|E_{ii}\rangle = a_{ii}$ donc :
$$A \in \{E_{i,i}\}^\perp \iff a_{ii} = 0\,,$$
et $\mathcal{D}_n(\mathbb{R})^\perp$ est l'ensemble des matrices de diagonale nulle.

■ Comment trouver et utiliser une base orthonormale ?

SF5.11 Utiliser la définition

Pour montrer que $(e_1,\dots,e_n)$ est une base orthonormale, il suffit de vérifier :

- que c'est une famille orthonormale : $\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \langle e_i\,|\,e_j\rangle = \delta_{ij}$;
- puis montrer soit qu'elle est génératrice, soit que son cardinal est égal à la dimension de E (si on la connaît).
  En effet, l'orthonormalité assure que la famille est libre.

Exemple

Soit E un espace préhilbertien réel, et $(e_1,\dots,e_n)$ une famille de vecteurs de E tels que : $\forall\, x \in E,\ \|x\|^2 = \sum_{i=1}^{n} \langle e_i\,|\,x\rangle^2$.

Montrer que cette famille est une base orthonormale de E dans chacun des deux cas suivants :

1) en supposant que $\|e_i\| \geqslant 1$ pour tout $i \in [\![1\,;n]\!]$;

2) en supposant que E est de dimension finie n.

1) *La relation donnée appliquée à chaque e_j pour $j \in [\![1\,;n]\!]$ donne :*

$$\|e_j\|^2 = \|e_j\|^4 + \sum_{i\neq j} \langle e_i \,|\, e_j\rangle^2 \,,$$

soit : $\displaystyle\sum_{i\neq j} \langle e_i \,|\, e_j\rangle^2 + \|e_j\|^2 \left(\|e_j\|^2 - 1\right) = 0\,.$

On a donc une somme de termes positifs (car $\|e_j\| \geqslant 1$ par hypothèse) qui est nulle, donc tous les termes sont nuls, ce qui implique que la famille (e_i) est orthonormale.

Montrons maintenant qu'il s'agit d'une base orthonormale de E (ici, on ne connaît pas la dimension de E !) Pour cela, considérons le sous-espace vectoriel F de E engendré par la famille orthonormale $(e_1,\ldots,e_n)$.
D'après une formule rappelée en **[S5.6]**, *la distance d'un vecteur x quelconque à F est donnée par :*

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{n} \langle e_i \,|\, x\rangle^2 \,,$$

et d'après l'hypothèse, cette quantité est nulle. Il en résulte que tout vecteur x appartient à F, c'est-à-dire $E = F$ donc $(e_1,\ldots,e_n)$ est bien une base orthonormale de E.

2) *L'idée utilisée ici a déjà été évoquée en* **[SF5.4]** *: puisque l'on connaît une relation sur la norme, on peut en déduire une relation sur le produit scalaire en utilisant les identités de polarisation.*
On suppose dans cette question $\dim E = n$. Soit donc $\mathscr{B}' = (e'_1,\ldots,e'_n)$ une base orthonormale de E. Pour $(i,j) \in [\![1\,;n]\!]^2$ on a :

$$\begin{aligned}
\delta_{ij} = \langle e'_i \,|\, e'_j\rangle &= \frac{1}{2}\left(\|e'_i + e'_j\|^2 - \|e'_i\|^2 - \|e'_j\|^2\right) \\
&= \frac{1}{2}\left(\sum_{k=1}^{n} \underbrace{\langle e_k \,|\, e'_i + e'_j\rangle^2}_{=\left(\langle e_k \,|\, e'_i\rangle + \langle e_k \,|\, e'_j\rangle\right)^2} - \sum_{k=1}^{n} \langle e_k \,|\, e'_i\rangle^2 - \sum_{k=1}^{n} \langle e_k \,|\, e'_j\rangle^2\right) \\
&= \sum_{k=1}^{n} \langle e_k \,|\, e'_i\rangle \langle e_k \,|\, e'_j\rangle \,.
\end{aligned}$$

Notons $M = (m_{ij})_{1\leqslant i,j\leqslant n}$ la matrice du système de vecteurs $(e_1,\ldots,e_n)$ dans $\mathscr{B}'$. $\mathscr{B}'$ étant orthonormale, $\langle e_k \,|\, e'_i\rangle$ est la i^e coordonnée de e_k dans $\mathscr{B}'$ c'est-à-dire m_{ik}.
La relation ci-dessus s'écrit donc :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \sum_{k=1}^{n} m_{ik} m_{jk} = \delta_{ij} \,,$$

soit encore $M\,{}^{\mathrm{t}}M = I_n$. La matrice M est donc une matrice orthogonale ; c'est la matrice de passage d'une base orthonormale à une base orthonormale, donc la famille $(e_1,\ldots,e_n)$ est bien une base orthonormale.

SF5.12 Utiliser le procédé d'orthonormalisation de Gram-Schmidt

Le procédé d'orthonormalisation de Gram-Schmidt a le gros avantage de donner une méthode constructive pour obtenir une base orthonormale. Cependant, sa mise en œuvre devient lourde dès que le nombre de vecteurs n'est pas petit. Heureusement, dans beaucoup de cas, il est surtout utilisé pour démontrer l'existence d'une base orthonormale adaptée à un problème donné ; il ne faut d'ailleurs pas oublier son interprétation matricielle, rappelée en **[S5.7]**.

Exemple

On munit $\mathbb{R}_3[X]$ du produit scalaire défini par :

$$\forall\, P,\, Q \in \mathbb{R}_3[X],\ \langle P \,|\, Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t\,.$$

Déterminer la base orthonormale obtenue à partir de la base canonique par le procédé de Gram-Schmidt.

Ce genre d'exercices est purement calculatoire et exige seulement de bien connaître les formules du cours rappelées en **[S5.7]**.

✓ *Une petite astuce à retenir pour diminuer (un peu) les calculs : on cherche d'abord une base orthogonale, puis on la normalise seulement à la fin. Cela permet d'éviter d'avoir des radicaux tout le long du calcul.*

Notons $a_0 = 1$, $a_1 = X$, $a_2 = X^2$ *et* $a_3 = X^3$ *la base canonique de* $\mathbb{R}_3[X]$. *On va construire une base orthogonale* (e'_0,e'_1,e'_2,e'_3) *à partir de celle-ci.*

- *On commence par prendre* $e'_0 = a_0$, *soit* $e'_0 = 1$.
- *Puis* $e'_1 = a_1 - \dfrac{\langle e'_0 \,|\, a_1\rangle}{\|e'_0\|^2} e'_0$. *Puisque* $\langle e'_0 \,|\, a_1\rangle = \displaystyle\int_{-1}^{1} t\,\mathrm{d}t = 0$, *on trouve* $e'_1 = a_1 = X$.
- *Ensuite* $e'_2 = a_2 - \dfrac{\langle e'_0 \,|\, a_2\rangle}{\|e'_0\|^2} e'_0 - \dfrac{\langle e'_1 \,|\, a_2\rangle}{\|e'_1\|^2} e'_1$. *On fait les calculs intermédiaires :*

  $$\|e'_0\|^2 = \int_{-1}^{1} \mathrm{d}t = 2\ ;\ \langle e'_0 \,|\, a_2\rangle = \int_{-1}^{1} t^2\,\mathrm{d}t = \frac{2}{3}\ ;\ \langle e'_1 \,|\, a_2\rangle = \int_{-1}^{1} t^3\,\mathrm{d}t = 0$$

  ce qui donne $e'_2 = X^2 - \dfrac{1}{3}$.
- *Enfin,* $e'_3 = a_3 - \dfrac{\langle e'_0 \,|\, a_3\rangle}{\|e'_0\|^2} e'_0 - \dfrac{\langle e'_1 \,|\, a_3\rangle}{\|e'_1\|^2} e'_1 - \dfrac{\langle e'_2 \,|\, a_3\rangle}{\|e'_2\|^2} e'_2$. *Or :*

  $$\langle e'_0 \,|\, a_3\rangle = \int_{-1}^{1} t^3\,\mathrm{d}t = 0\ ;\ \langle e'_1 \,|\, a_3\rangle = \int_{-1}^{1} t^4\,\mathrm{d}t = \frac{2}{5}\ ;\ \langle e'_2 \,|\, a_3\rangle = 0\ ;$$

  $$\text{et}\quad \|e'_1\|^2 = \int_{-1}^{1} t^2\,\mathrm{d}t = \frac{2}{3}\,.$$

  On obtient donc : $e'_3 = X^3 - \dfrac{3}{5}X$.
- *Il reste ensuite à normer les vecteurs obtenus ; si l'on pose* $e_i = \dfrac{e'_i}{\|e'_i\|}$ *on trouve après calculs :*

$$e_0 = \frac{\sqrt{2}}{2}\ ;\ e_1 = \frac{\sqrt{6}}{2}X\ ;\ e_2 = \frac{3\sqrt{10}}{4}\left(X^2 - \frac{1}{3}\right)\ ;\ e_3 = \frac{5\sqrt{14}}{4}\left(X^3 - \frac{3}{5}X\right).$$

L'exemple qui suit est assez délicat, mais montre bien comment on peut utiliser le procédé de Schmidt dans un cadre plus théorique.

Exemple

1) Soit $A \in \mathrm{GL}_n(\mathbb{R})$. Montrer qu'il existe une matrice orthogonale $Q \in \mathcal{O}_n(\mathbb{R})$ et une matrice triangulaire supérieure $R \in \mathcal{T}_n^+(\mathbb{R})$ à éléments diagonaux positifs telles que $A = QR$.

2) Prouver l'unicité de cette décomposition.

3) Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$, dont on note $C_1, \ldots, C_n$ les colonnes. Prouver que :

$$|\det A| \leqslant \|C_1\| \cdots \|C_n\| \ ,$$

où $\| \ \|$ désigne la norme euclidienne usuelle dans $\mathbb{R}^n$; préciser les cas d'égalité.

On notera $\mathscr{B}$ la base canonique de $\mathbb{R}^n$, et on suppose $\mathbb{R}^n$ muni de sa structure euclidienne canonique.
Comme d'habitude, on identifiera le vecteur $x \in \mathbb{R}^n$ et la matrice colonne $X \in \mathcal{M}_{n,1}(\mathbb{R})$ formée des coordonnées de x dans la base $\mathscr{B}$.

1) *Puisque A est inversible, ses vecteurs colonnes $C_1, \ldots, C_n$ forment une base de $\mathbb{R}^n$, que nous noterons $\mathscr{B}'$; ainsi, A est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.*
Notons alors $\mathscr{B}''$ la base obtenue à partir de $\mathscr{B}'$ par le procédé de Schmidt ; on a expliqué en **[S5.7]** *pourquoi la matrice de passage R de $\mathscr{B}''$ à $\mathscr{B}'$ est triangulaire supérieure à éléments diagonaux strictement positifs.*
Les formules de changement de base rappelées en **[S3.10]** *donnent :*

$$P_{\mathscr{B}}^{\mathscr{B}'} = P_{\mathscr{B}}^{\mathscr{B}''} \times P_{\mathscr{B}''}^{\mathscr{B}'}$$

ce qui s'écrit aussi $A = QR$, en notant Q la matrice de passage de $\mathscr{B}$ à $\mathscr{B}''$. Puisque les bases $\mathscr{B}$ et $\mathscr{B}''$ sont orthonormales, Q est une matrice orthogonale, ce qui établit le résultat voulu.

2) *Supposons qu'il existe deux décompositions : $A = QR = Q'R'$, avec Q, Q' orthogonales et R, R' triangulaires supérieures à éléments diagonaux strictement positifs.*
Alors ${Q'}^{-1}Q = R'R^{-1}$; la matrice $M = {Q'}^{-1}Q$ est donc orthogonale (car $\mathcal{O}_n(\mathbb{R})$ est un groupe pour la multiplication) et aussi triangulaire supérieure à éléments diagonaux strictement positifs (cf. propriétés des matrices triangulaires supérieures en **[S3.5]**). *Puisque $M^{-1} = {}^tM$, elle est aussi triangulaire inférieure, donc diagonale. La même relation montre alors que ses éléments diagonaux sont égaux à leurs inverses, donc valent ± 1, et puisqu'elle est à éléments diagonaux positifs, on a finalement $M = I_n$ d'où $Q = Q'$ et $R = R'$, ce qui prouve l'unicité.*

3) *La relation demandée est immédiate si* $\det A = 0$.
Sinon, A est inversible et on peut appliquer le résultat de la première question. Avec les mêmes notations, puisque le déterminant d'une matrice orthogonale vaut ± 1 et que les termes diagonaux de R sont positifs, on aura :

$$|\det A| = |\det Q| \, |\det R| = \det R \, .$$

Or puisque R est la matrice de passage de $\mathscr{B}'' = (e''_1,\dots,e''_n)$, orthonormale, à $\mathscr{B}' = (C_1,\dots,C_n)$, on a $r_{ij} = \langle C_j \,|\, e''_i\rangle$ et en particulier $r_{ii} = \langle C_i \,|\, e''_i\rangle$ d'où d'après l'inégalité de Cauchy-Schwarz $r_{ii} \leqslant \|C_i\|\,\|e''_i\| = \|C_i\|$. Puisque $\det R = \prod_{i=1}^{n} r_{ii}$, on obtient le résultat demandé.

Nous vous laissons le soin de vérifier, en utilisant le cas d'égalité dans l'inégalité de Cauchy-Schwarz, qu'il y a égalité si et seulement si la famille $(C_1,\dots,C_n)$ est orthogonale.

■ Comment montrer qu'une projection est orthogonale ?

SF5.13 Revenir à la définition

Pour montrer qu'un endomorphisme p d'un espace préhilbertien E est une projection orthogonale en utilisant les définitions, on peut procéder ainsi :

- d'abord on montre qu'il s'agit bien d'une projection ; pour cela, on peut vérifier $p \circ p = p$, ou bien montrer que p est diagonalisable de spectre inclus dans $\{0,1\}$ (car alors $E = \operatorname{Ker} p \oplus E_1(p)$, et on retrouve la définition géométrique d'une projection) ;
- puis on vérifie ensuite :
  - que pour tout $x \in E$, $p(x)$ et $x - p(x)$ sont orthogonaux ;
  - ou bien que : $\forall\,(x,y) \in \operatorname{Ker} p \times \operatorname{Im} p,\ \langle x \,|\, y\rangle = 0$.

Exemple

Soit p un projecteur d'un espace préhilbertien E.

Montrer que p est orthogonal si et seulement si : $\forall\, x \in E,\ \langle x \,|\, p(x)\rangle \geqslant 0$.

- *Supposons d'abord que p est un projecteur orthogonal. Alors pour tout $x \in E$, $x - p(x) \perp p(x)$ donc :*

$$\begin{aligned}\langle p(x) \,|\, x\rangle &= \langle p(x) \,|\, x - p(x) + p(x)\rangle \\ &= \underbrace{\langle p(x) \,|\, x - p(x)\rangle}_{=0} + \langle p(x) \,|\, p(x)\rangle = \|p(x)\|^2 ,\end{aligned}$$

  d'où $\langle p(x) \,|\, x\rangle \geqslant 0$.
- *Réciproquement, supposons la propriété de l'énoncé vérifiée ; nous allons démontrer que $\operatorname{Im} p$ et $\operatorname{Ker} p$ sont orthogonaux.*
  Pour cela, soient $x \in \operatorname{Im} p$ et $y \in \operatorname{Ker} p$. Alors pour tout réel λ,

$$\langle p(x + \lambda y) \,|\, x + \lambda y\rangle \geqslant 0 ,$$

  ce qui donne, puisque $p(x) = x$ (p étant un projecteur, on rappelle que $\operatorname{Im} p = E_1(p)$) et $p(y) = 0$:

$$\forall\, \lambda \in \mathbb{R},\ \langle x \,|\, x + \lambda y\rangle = \|x\|^2 + \lambda\,\langle x \,|\, y\rangle \geqslant 0 .$$

  Or si $\langle x \,|\, y\rangle$ était non nul, l'expression $\|x\|^2 + \lambda\,\langle x \,|\, y\rangle$ changerait de signe pour $\lambda = -\dfrac{\|x\|^2}{\langle x \,|\, y\rangle}$, ce qui est contradictoire.

On a donc forcément $\langle x\,|\,y\rangle = 0$, *ce qui prouve* $\operatorname{Ker} p \perp \operatorname{Im} p$ *et montre que* p *est un projecteur orthogonal.*

SF5.14 Utiliser la caractérisation d'une projection orthogonale

Il s'agit d'utiliser ici l'une des deux caractérisations suivantes :

- un projecteur p est orthogonal si et seulement si il diminue la norme :
$$\forall\, x \in E,\ \|p(x)\| \leqslant \|x\| \ .$$
- un projecteur p est orthogonal si et seulement si c'est un endomorphisme symétrique, c'est-à-dire :
$$\forall\, (x,y) \in E^2,\ \langle p(x)\,|\,y\rangle = \langle x\,|\,p(y)\rangle \ .$$
Dans le cas où E est de dimension finie, cela équivaut à dire que la matrice de p dans une base orthonormale est symétrique.

> ✓ Avant d'utiliser l'une de ces deux caractérisations, il ne faut pas oublier de montrer que p est un projecteur, c'est-à-dire $p^2 = p$!

Exemple

Soient p et q deux projecteurs orthogonaux d'un espace vectoriel préhilbertien E. Montrer que les deux propriétés suivantes sont équivalentes :

(i) $p \circ q = q \circ p$;

(ii) $p \circ q$ est un projecteur.

- $(i) \Longrightarrow (ii)$ *est immédiat puisque, si* (i) *est vérifié alors :*
$$(p \circ q)^2 = p \circ q \circ p \circ q = p \circ p \circ q \circ q = p \circ q \,.$$
Pour cette partie, on n'utilise pas le fait que p *et* q *sont orthogonaux.*
- *Réciproquement, supposons que* $p \circ q$ *est un projecteur. On a alors :*
$$\forall\, x \in E,\ \|p \circ q(x)\| = \|p(q(x))\| \leqslant \|q(x)\| \leqslant \|x\|$$
puisque p *et* q *sont des projecteurs orthogonaux.*
Il en résulte que $p \circ q$ *est un projecteur orthogonal ; on aura donc, pour tous* x, $y \in E$, *en utilisant encore le fait que* p *et* q *sont des projecteurs orthogonaux donc des endomorphismes symétriques :*
$$\langle p \circ q(x)\,|\,y\rangle = \langle x\,|\,p \circ q(y)\rangle = \langle p(x)\,|\,q(y)\rangle = \langle q \circ p(x)\,|\,y\rangle$$
soit $\big\langle (p \circ q - q \circ p)(x)\,\big|\,y\big\rangle = 0$.
Cela signifie que le vecteur $(p \circ q - q \circ p)(x)$ *est orthogonal à tout l'espace, donc est nul, et cela étant vrai pour tout* x, *cela prouve que* $p \circ q = q \circ p$.

> ✓ *L'implication* $(ii) \Longrightarrow (i)$ *peut être fausse si l'on ne suppose plus* p *et* q *orthogonaux, comme le montre l'exemple des projecteurs de* $\mathbb{R}^2$ *de matrices respectives* $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ *et* $\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ *(vérifiez !).*

Exemple

Montrer que la matrice $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ avec $a_{ij} = \dfrac{1}{n}$ pour tous i, j est celle d'un projecteur orthogonal de $\mathbb{R}^n$.

Comme d'habitude, on suppose $\mathbb{R}^n$ muni de sa structure euclidienne canonique, pour laquelle la base canonique est orthonormale.

- *D'après la formule du produit matriciel,*
$$\forall\,(i,j)\in[\![1\,;n]\!]^2,\ (A^2)_{ij}=\sum_{k=1}^{n}\frac{1}{n^2}=\frac{1}{n}=a_{ij}$$
donc $A^2=A$; A est bien une matrice de projection.
- *Puisque $A={}^{\mathrm{t}}A$, l'endomorphisme canoniquement associé à A dans la base canonique est symétrique ; c'est donc bien une projection orthogonale.*
- *Enfin, toutes les colonnes de A étant égales, $\operatorname{Im}A$ est la droite vectorielle de $\mathbb{R}^n$ engendrée par le vecteur $(1,\dots,1)$, et A est donc la matrice de la projection orthogonale sur cette droite.*

■ Comment trouver l'expression d'une projection ou d'une symétrie orthogonale ?

Si p_F est la projection orthogonale sur un sous-espace vectoriel F de E et si s_F est la symétrie orthogonale par rapport à F, on sait que l'on a :
$$s_F=2p_F-\mathrm{Id}_E\,.$$
Donc pour trouver l'expression de s_F, il suffira de trouver celle de p_F, ce qui est en général un peu plus simple.

SF5.15 Utiliser la formule

Si l'on connaît une base *orthogonale* $(e_1,\dots,e_p)$ de F, le projeté orthogonal d'un vecteur x sur F est donné par :
$$p_F(x)=\sum_{i=1}^{p}\frac{\langle e_i\,|\,x\rangle}{\|e_i\|^2}\,e_i\,.$$
Cette formule permet donc d'obtenir directement $p_F(x)$. En dimension finie, elle permet de calculer les coordonnées de $p_F(x)$ dans une base donnée de E, donc d'obtenir la matrice de p_F dans cette base.

✓ Si l'on connaît l'expression de la projection orthogonale p_F sur F, on connaît aussi celle de la projection orthogonale $p_{F^\perp}$ sur $F^\perp$ puisque $p_F+p_{F^\perp}=\mathrm{Id}_E$.
Pour minimiser le nombre de calculs si l'on utilise la formule ci-dessus, on aura donc intérêt à déterminer la projection orthogonale sur celui des espaces F et $F^\perp$ qui a la plus petite dimension.

Exemple

Dans $\mathbb{R}^4$ muni de sa structure euclidienne canonique, déterminer la matrice dans la base canonique de la symétrie orthogonale par rapport au sous-espace vectoriel P dont un système d'équations est :
$$x+y+z+t=0\quad\text{et}\quad x+2y+3z+4t=0\,.$$

- *On remarque déjà que P est un plan vectoriel, puisque c'est l'intersection de deux hyperplans distincts de $\mathbb{R}^4$.*
- *On peut chercher une base de P par la méthode habituelle de résolution d'un système :*

$$u=(x,y,z,t)\in P \iff \left\{\begin{array}{rcl} x+y &=& -z-t \\ x+2y &=& -3z-4t \end{array}\right. \iff \left\{\begin{array}{rcl} x &=& z+2t \\ y &=& -2z-3t \end{array}\right.$$

donc : $u\in P \iff u=(z+2t,-2z-3t,z,t)=z(1,-2,1,0)+t(2,-3,0,1)$, ce qui montre qu'une base de P est formée des vecteurs $a_1=(1,-2,1,0)$ et $a_2=(2,-3,0,1)$. Le procédé de Schmidt conduit à considérer le vecteur

$$a_2-\frac{\langle a_2\,|\,a_1\rangle}{\|a_1\|^2}a_1=a_2-\frac{8}{6}a_1=\left(\frac{2}{3},-\frac{1}{3},-\frac{4}{3},1\right)$$

et ainsi les vecteurs $e_1=a_1=(1,-2,1,0)$ et $e_2=(2,-1,-4,3)$ forment une base orthogonale de P.

- *Si l'on note p la projection orthogonale sur P on aura donc, pour $u=(x,y,z,t)$:*

$$\begin{aligned} p(u) &= \frac{\langle e_1\,|\,u\rangle}{\|e_1\|^2}e_1+\frac{\langle e_2\,|\,u\rangle}{\|e_2\|^2}e_2 \\ &= \frac{x-2y+z}{6}(1,-2,1,0)+\frac{2x-y-4z+3t}{30}(2,-1,-4,3) \end{aligned}$$

donc, en notant (x',y',z',t') les coordonnées de $p(u)$, on a après quelques calculs :

$$\left\{\begin{array}{lll} x' = \dfrac{3x-4y-z+2t}{10} & & z' = \dfrac{-x-2y+7z-4t}{10} \\ y' = \dfrac{-4x+7y-2z-t}{10} & & t' = \dfrac{2x-y-4z+3t}{10} \end{array}\right.$$

ce qui donne la matrice de p dans la base canonique :

$$M(p)=\frac{1}{10}\begin{pmatrix} 3 & -4 & -1 & 2 \\ -4 & 7 & -2 & -1 \\ -1 & -2 & 7 & -4 \\ 2 & -1 & -4 & 3 \end{pmatrix}.$$

> ✓ *On constate que la matrice obtenue est symétrique : cela est normal puisqu'une projection orthogonale est un endomorphisme symétrique, et permet de vérifier les calculs faits.*

- *Enfin, si s est la symétrie orthogonale par rapport à P, puisque $s=2p-\mathrm{Id}_{\mathbb{R}^4}$ on a :*

$$M(s)=2M(p)-I_4=\frac{1}{5}\begin{pmatrix} -2 & -4 & -1 & 2 \\ -4 & 2 & -2 & -1 \\ -1 & -2 & 2 & -4 \\ 2 & -1 & -4 & -2 \end{pmatrix}.$$

SF5.16 Revenir à la définition d'une projection ou d'une symétrie

Le gros défaut de la formule précédente est qu'il faut connaître une base orthogonale de F, ce qui peut conduire à des calculs pénibles.
Une méthode bien plus rapide en général consiste simplement à revenir à la définition d'une projection.

Si p est une projection sur F parallèlement à un sous-espace vectoriel supplémentaire G, on sait qu'elle est caractérisée par :

$$y = p_F(x) \iff y \in F \text{ et } x - y \in G\,.$$

Dans le cas de la projection orthogonale p_F cela devient :

$$y = p_F(x) \iff y \in F \text{ et } x - y \in F^\perp\,.$$

De même, si on note s_F la symétrie orthogonale par rapport à F on peut écrire directement :

$$y = s_F(x) \iff x + y \in F \text{ et } x - y \in F^\perp\,.$$

Cette caractérisation permet éventuellement d'utiliser une base quelconque, pas forcément orthogonale, de F ou de $F^\perp$.

✓ Nous vous suggérons de reprendre l'exemple précédent en utilisant cette méthode, afin de le traiter de façon plus élémentaire. En reprenant les mêmes notations, le vecteur $s(u)$ est ici caractérisé par $u + s(u) \in F$ et $\langle u - s(u) \,|\, a_1\rangle = \langle u - s(u) \,|\, a_2\rangle = 0$. Ces équations donnent directement un système permettant d'obtenir les coordonnées de $s(u)$.

Exemple

On se place dans l'espace vectoriel $E = \mathbb{R}[X]$ que l'on munit du produits scalaire

$$\langle P \,|\, Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t\,.$$

Déterminer le projeté orthogonal de X^2 sur $\mathbb{R}_1[X]$.

Notons P ce projeté.
Puisque $P \in \mathbb{R}_1[X]$, il existe des réels a et b tels que $P = aX + b$.
On doit aussi avoir $X^2 - P \in \mathbb{R}_1[X]^\perp$ et puisque $\mathbb{R}_1[X] = \mathrm{Vect}(1,X)$, cela équivaut à :

$$\big\langle X^2 - P \,\big|\, 1\big\rangle = \big\langle X^2 - P \,\big|\, X\big\rangle = 0\,.$$

En appliquant la définition du produit scalaire cela donne :

$$\int_0^1 (t^2 - at - b)\,\mathrm{d}t = \int_0^1 t(t^2 - at - b)\,\mathrm{d}t = 0$$

d'où les équations : $\frac{1}{3} - \frac{a}{2} - b = 0$ *et* $\frac{1}{4} - \frac{a}{3} - \frac{b}{2} = 0$, *qui donnent :* $a = 1$ *et* $b = -\frac{1}{6}$.

Le projeté orthogonal cherché est donc $P = X - \frac{1}{6}$.

■ Comment utiliser les projections orthogonales ?

Une projection orthogonale est avant tout une projection, et elle en possède donc toutes les propriétés. Mais ce qui distingue une projection orthogonale parmi les projections, c'est qu'il s'agit de la seule qui réalise la distance entre un vecteur x et un sous-espace vectoriel de dimension finie F :

$$y \in F \text{ et } d(x,F) = \|x - y\| \iff y = p_F(x)\,.$$

SF5.17 Résoudre un problème de minimum

Certains énoncés d'analyse concernant la recherche d'un minimum peuvent parfois être traités en utilisant des méthodes d'algèbre linéaire. Pour cela, il faut interpréter le problème posé en termes de distance et de minimum de cette distance, donc il faut :

- déterminer l'espace vectoriel dans lequel on travaille ;
- préciser le produit scalaire utilisé ;
- préciser le sous-espace vectoriel sur lequel on projette ;
- puis enfin calculer explicitement soit le projeté orthogonal, soit seulement la distance (selon l'énoncé).

Exemple

Calculer $\displaystyle\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^{+\infty} (t^3 - at^2 - bt - c)^2 \mathrm{e}^{-t}\,\mathrm{d}t$.

- *On commence déjà par remarquer rapidement que l'intégrale écrite a bien un sens : en effet, la fonction* $t \mapsto (t^3 - at^2 - bt - c)^2\mathrm{e}^{-t}$ *est continue sur* $\mathbb{R}_+$ *et elle est négligeable devant* $t \mapsto \dfrac{1}{t^2}$ *au voisinage de* $+\infty$.
- *L'énoncé suggère de considérer l'espace vectoriel* $\mathbb{R}_3[X]$ *des polynômes de degré* $\leqslant 3$, *muni du produit scalaire :* $(P,Q) \mapsto \displaystyle\int_0^{+\infty} P(t)Q(t)\mathrm{e}^{-t}\,\mathrm{d}t$.

  Le fait qu'il s'agit bien d'un produit scalaire sur $\mathbb{R}_3[X]$ *se démontre aisément (voir aussi l'exercice* **5.5** *ci-après).*

  ✓ *Pour pouvoir reconnaître la situation adaptée à l'exercice, il faut évidemment connaître un certain nombre de produit scalaires « usuels ».*
- *Si l'on considère le polynôme* $Q_{a,b,c} = X^3 - aX^2 - bX - c$, *on a donc, pour ce produit scalaire :*

$$\int_0^{+\infty} (t^3 - at^2 - bt - c)^2 \mathrm{e}^{-t}\,\mathrm{d}t = \|Q_{a,b,c}\|^2\,.$$

  Or $Q_{a,b,c} = P_0 - P_{a,b,c}$ *avec* $P_0 = X^3$ *et* $P_{a,b,c} = aX^2 + bX + c$; *lorsque* (a,b,c) *décrit* $\mathbb{R}^3$, $P_{a,b,c}$ *décrit le sous-espace vectoriel* $F = \mathbb{R}_2[X]$.
  Le problème posé revient donc à chercher :

$$\inf_{(a,b,c)\in\mathbb{R}^3} \|P_0 - P_{a,b,c}\|^2 = \inf_{P\in F} \|P_0 - P\|^2\,,$$

  c'est-à-dire la distance (au carré) entre P_0 *et* F.

- *Pour cela, cherchons le polynôme $P = aX^2 + bX + c$ qui est le projeté orthogonal de P_0 sur F. Il doit donc vérifier :*

$$\forall\, k \in \{0,1,2\}\,,\ \langle P_0 - P \,|\, X^k \rangle = 0 \text{ soit } \int_0^{+\infty} t^k(t^3 - at^2 - bt - c)\,\mathrm{d}t = 0\,.$$

Sachant que $\displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\,\mathrm{d}t = n!$ (résultat classique qui se démontre par récurrence sur n), on obtient le système :

$$6 - 2a - b - c = 0\ ;\ 24 - 6a - 2b - c = 0\ ;\ 120 - 24a - 6b - 2c = 0$$

qui se résout en $a = 9$, $b = -18$ et $c = 6$.
Cela signifie que le polynôme $P = 9X^2 - 18X + 6$ est le projeté orthogonal de P_0 sur F, mais ne répond pas tout à fait à la question, qui est de calculer $d^2 = \|P_0 - P\|^2$. Bien sûr on pourrait remplacer P et P_0 par leurs expressions, mais il y a une petite astuce qui minimise le nombre de calculs.
En effet, $\|P_0 - P\|^2 = \langle P_0 - P \,|\, P_0 - P\rangle = \langle P_0 \,|\, P_0 - P\rangle - \langle P \,|\, P_0 - P\rangle$ et puisque $P_0 - P$ appartient à $F^\perp$ et $P \in F$, ce dernier produit scalaire est nul.
Il reste donc :

$$\begin{aligned} d^2 = \|P_0 - P\|^2 = \langle P_0 \,|\, P_0 - P\rangle &= \int_0^{+\infty} t^3(t^3 - at^2 - bt - c)\,\mathrm{d}t \\ &= 720 - 120a - 24b - 6c \end{aligned}$$

et en remplaçant a,b et c par leurs valeurs, on trouve que le minimum cherché vaut $d^2 = 36$.

À vous de jouer !

Exercice 5.1 **SF5.3**

Soit $n \in \mathbb{N}^*$, et $x_1, \ldots, x_n$ des réels strictement positifs tels que : $\displaystyle\sum_{i=1}^{n} x_i = 1$.

Montrer que : $\displaystyle\sum_{i=1}^{n} \frac{1}{x_i} \geqslant n^2$, et préciser le cas d'égalité.

Exercice 5.2 **SF5.15** **SF5.17**

On se place ici dans $\mathcal{M}_n(\mathbb{R})$ munit du produit scalaire canonique.

1) Montrer que l'ensemble H des matrices de trace nulle est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ et en donner la dimension.

2) Soit J la matrice dont tous les coefficients sont égaux à 1. Calculer $d(J,H)$.

Exercice 5.3 SF5.12 SF5.15 SF5.16 SF5.17

Soient $\big((x_1,\dots,x_n),(y_1,\dots,y_n)\big) \in (\mathbb{R}^n)^2$. On suppose que les x_i ne sont pas tous égaux entre eux.

Interpréter $\varphi(a,b) = \sum_{i=1}^{n}(y_i - ax_i - b)^2$ à l'aide de la distance d'un vecteur de $\mathbb{R}^n$ à un vecteur d'un sous-espace de $\mathbb{R}^n$.

Montrer que φ admet un minimum absolu et déterminer le couple (a,b) correspondant en fonction de $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ et $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$.

Exercice 5.4 SF5.5

Soit C une partie convexe compacte non vide d'un espace préhilbertien réel E.

1) Soit $x \in E$. Montrer qu'il existe un unique $x_0 \in C$ tel que $d(x,C) = \|x - x_0\|$. On note $x_0 = p_C(x)$; x_0 s'appelle le projeté de x sur C.

2) Montrer que pour tout $x \in E$ on a :

$$\forall\, y \in C,\ \langle x - p_C(x) \,|\, y - p_C(x)\rangle \leqslant 0\,.$$

Exercice 5.5 SF5.1 SF5.11

On note E l'ensemble des fonctions $f \in \mathscr{C}([0\,;+\infty[,\mathbb{R})$ telles que $\displaystyle\int_0^{+\infty} f^2(t)\mathrm{e}^{-t}\,\mathrm{d}t$ existe, et pour tous f, $g \in E$, on pose $\langle f \,|\, g\rangle = \displaystyle\int_0^{+\infty} f(t)g(t)\mathrm{e}^{-t}\,\mathrm{d}t$.

1) Montrer que E est un sous-espace vectoriel de $\mathscr{C}([0\,;+\infty[,\mathbb{R})$ qui contient l'espace vectoriel des fonctions polynômes, et qu'on définit ainsi un produit scalaire sur E.

2) Pour $x \in [0\,;+\infty[$, on pose $L_n(x) = \dfrac{1}{n!}\dfrac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n\mathrm{e}^{-x})\,\mathrm{e}^{x}$.

Montrer que la famille $(L_n)_{n\in\mathbb{N}}$ est une famille orthonormale de polynômes.

3) Pour tout $\alpha > 0$, on pose $f_\alpha : \left\{\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \mathrm{e}^{-\alpha x}\,. \end{array}\right.$

Calculer $\|f_\alpha\|^2$, $\langle f_\alpha \,|\, L_n\rangle$ et $\displaystyle\sum_{n=0}^{+\infty} \langle f_\alpha \,|\, L_n\rangle^2$. Interprétation ?

Exercice 5.6 SF5.6 SF5.9

Soit $A \in \mathcal{M}_n(\mathbb{R})$, telle que ${}^{\mathrm{t}}AA = A^2$.

1) Montrer que $\mathbb{R}^n = \mathrm{Ker}(A) \overset{\perp}{\oplus} \mathrm{Im}(A)$ (somme directe orthogonale).

2) Montrer que A est symétrique.

Exercice 5.7 **SF3.1** **SF3.35** **SF3.37** **SF5.13** **SF5.17**

Soit E un espace préhilbertien réel. Pour toute famille $(v_1,\dots,v_n)$ de vecteurs de E, on considère la matrice :

$$G(v_1,\dots,v_n) = (\langle v_i \,|\, v_j\rangle)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}} .$$

Soit $(v_1,\dots,v_n)$ une famille de vecteurs de E, et F un sous-espace vectoriel de dimension n contenant cette famille. Soit $\mathscr{B}$ une base orthonormale de F et A la matrice dans $\mathscr{B}$ de la famille $(v_1,\dots,v_n)$.

1) Montrer que : $G(v_1,\dots,v_n) = {}^{\mathrm{t}}AA$.

2) En déduire que : $\det G(v_1,\dots,v_n) \geqslant 0$ et que

$$(v_1,\dots,v_n) \text{ libre} \iff \det G(v_1,\dots,v_n) \neq 0 .$$

3) Soit p la projection orthogonale sur $\mathrm{Vect}(v_1,\dots,v_n)$. Montrer que :

$$\forall\, x \in E,\ \det G(x,v_1,\dots,v_n) = \langle x \,|\, x - p(x)\rangle\, G(v_1,\dots,v_n) .$$

4) On suppose que la famille $(v_1,\dots,v_n)$ est libre. Montrer que :

$$\forall\, x \in E,\ d(x,F)^2 = \frac{\det G(x,v_1,\dots,v_n)}{\det G(v_1,\dots,v_n)} .$$

Exercice 5.8 **SF5.14**

Soit E un espace préhilbertien réel, et $(p_n)_{n\in\mathbb{N}}$ une suite de projecteurs orthogonaux. On suppose que cette suite converge simplement, c'est-à-dire qu'il existe une application $u\colon E \to E$ telle que :

$$\forall\, x \in E,\ \lim_{n\to+\infty} p_n(x) = u(x) .$$

Montrer que u est un projecteur orthogonal.

Exercice 5.9 **SF5.16**

Soient a et b deux vecteurs distincts et de même norme d'un espace vectoriel préhilbertien E.

Montrer qu'il existe une et une seule réflexion s (symétrie orthogonale par rapport à un hyperplan) telle que $s(a) = b$.

Thème 6 - Endomorphismes des espaces euclidiens

Dans tout ce chapitre, E désigne un espace euclidien (= préhilbertien réel de dimension finie), muni d'un produit scalaire noté $\langle \cdot | \cdot \rangle$.

■ Comment montrer qu'un endomorphisme est symétrique ?

SF6.1 Utiliser la définition

Il s'agit tout simplement de faire un calcul de produit scalaire :

$$u \in \mathscr{L}(E) \text{ est symétrique } \iff \forall\,(x,y) \in E^2,\ \langle u(x)\,|\,y\rangle = \langle x\,|\,u(y)\rangle\,.$$

Exemple

Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, montrer que l'endomorphisme :

$$u\colon M \mapsto \alpha M + \mathrm{tr}(M)\mathrm{I}_n \quad (\alpha \in \mathbb{R})$$

est symétrique.

- *Le fait que u est bien un endomorphisme de $\mathcal{M}_n(\mathbb{R})$ est clair.*
- *Pour M, $N \in \mathcal{M}_n(\mathbb{R})$ on calcule :*

  $$\langle u(M)\,|\,N\rangle = \langle \alpha M + \mathrm{tr}(M)\mathrm{I}_n\,|\,N\rangle = \alpha\,\langle M\,|\,N\rangle + \mathrm{tr}(M)\,\langle \mathrm{I}_n\,|\,N\rangle\ ,$$

  et puisque $\langle \mathrm{I}_n\,|\,N\rangle = \mathrm{tr}(N)$ (par définition du produit scalaire $\langle A\,|\,B\rangle = \mathrm{tr}({}^tAB)$), on a :

  $$\langle u(M)\,|\,N\rangle = \alpha\,\langle M\,|\,N\rangle + \mathrm{tr}(M)\,\mathrm{tr}(N)\,.$$

  Le résultat étant symétrique en M et N, u est bien un endomorphisme symétrique pour le produit scalaire considéré.

SF6.2 Utiliser les matrices

Il s'agit d'utiliser la propriété :

> u est symétrique si et seulement si sa matrice dans une *base orthonormale* est symétrique.

Cette propriété peut être très commode ; néanmoins, comme dans tous les problèmes d'algèbre linéaire, le recours au calcul matriciel ne doit pas être systématique.

Exemple

Caractériser l'endomorphisme p de $\mathbb{R}^3$ dont la matrice dans la base canonique est :

$$A = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\,.$$

Bien sûr, on munit $\mathbb{R}^3$ de son produit scalaire canonique, pour lequel la base canonique est une base orthonormale.
Un calcul rapide montre que $A^2 = A$; donc $p^2 = p$, et p est un projecteur.
Puisque A est symétrique, p est un endomorphisme symétrique ; c'est donc une projection orthogonale (voir la caractérisation en **[S5.9]***).*
La résolution du système $A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ (que nous ne détaillons pas) donne $x = y = z$; ainsi le noyau de p est la droite de base le vecteur $(1,1,1)$. On en déduit que l'image de p est le plan orthogonal à cette droite, et on peut donc conclure :

p est la projection orthogonale sur le plan d'équation $x + y + z = 0$.

✓ *On sait que pour une projection on a $\operatorname{tr} p = \operatorname{rg} p = \dim \operatorname{Im} p$. Ici, on pouvait vérifier que l'on a bien $\operatorname{tr} A = 2$. Ce genre de vérification, très rapide à faire, peut parfois permettre de détecter une erreur de calcul.*

Exemple

Soient f, g deux endomorphismes symétriques d'un espace vectoriel euclidien E.

Montrer que $f \circ g$ est symétrique si et seulement si $f \circ g = g \circ f$.

Notons A et B les matrices respectives de f et g dans une base orthonormale $\mathscr{B}$. Par hypothèse, A et B sont symétriques.
La matrice dans la même base $\mathscr{B}$ de $f \circ g$ est égale à AB. On a donc :

$$f \circ g \text{ symétrique} \iff {}^{\mathrm{t}}(AB) = AB \iff {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A = AB \iff BA = AB$$

et cela établit l'équivalence demandée.

■ Comment utiliser les endomorphismes symétriques ?

SF6.3 Utiliser la définition avec le produit scalaire

Exemple

Soient p et q deux projecteurs orthogonaux d'un espace euclidien E.

Montrer que $\operatorname{Sp}(p \circ q) \subset [0\,;1]$.

Rappelons encore une fois qu'un projecteur orthogonal est aussi un endomorphisme symétrique (cf. **[S5.9]***).*
Soit λ une valeur propre de $p \circ q$: il existe $x \neq 0$ tel que $p \circ q(x) = \lambda x$.

– *Si $\lambda = 0$, le résultat est acquis.*

– *Sinon, puisque $x = p \circ q\left(\frac{x}{\lambda}\right)$, on aura $x \in \operatorname{Im} p$ donc $p(x) = x$. D'où*

$$\lambda \|x\|^2 = \langle p \circ q(x) \,|\, x \rangle \underbrace{=}_{\substack{car\ p \\ sym.}} \langle q(x) \,|\, p(x) \rangle = \langle q(x) \,|\, x \rangle \ .$$

q étant un projecteur orthogonal, on a :

$$\begin{aligned}\langle q(x)\,|\,x\rangle &= \langle q(x)\,|\,q(x)+x-q(x)\rangle \\ &= \langle q(x)\,|\,q(x)\rangle + \underbrace{\langle q(x)\,|\,x-q(x)\rangle}_{=0} = \|q(x)\|^2\ ,\end{aligned}$$

et finalement : $\lambda\|x\|^2 = \|q(x)\|^2$.
x étant non nul on en déduit $\lambda \geqslant 0$ et, puisque $\|q(x)\| \leqslant \|x\|$ (cf. encore **[S5.9]***) , on en déduit $\lambda\|x\|^2 \leqslant \|x\|^2$ puis $\lambda \leqslant 1$, ce qui achève la démonstration.*

SF6.4 Utiliser le calcul matriciel

Il s'agit ici d'utiliser l'écriture matricielle du produit scalaire. Citons deux exemples de calculs particulièrement fréquents, et qu'il faut savoir refaire à chaque fois.

- Si l'on se donne une base orthonormale $\mathscr{B}$, deux vecteurs x et y dont les matrices colonnes des coordonnées dans $\mathscr{B}$ sont X et Y respectivement, et si u est un endomorphisme de E de matrice A dans $\mathscr{B}$, on remarque que :

$$\langle u(x)\,|\,y\rangle = {}^t(AX)Y = {}^tX{}^tAY\ .$$

  Lorsque u est symétrique, on a ${}^tA = A$ et alors :

$$\langle u(x)\,|\,y\rangle = {}^tXAY = \langle x\,|\,u(y)\rangle\ .$$

  Il peut d'ailleurs être intéressant de se rappeler (voir aussi **[SF5.2]**) que, avec les notations habituelles :

$${}^tXAY = \sum_{1\leqslant i,j\leqslant n} a_{ij}x_iy_j\ .$$

- Soit $u \in \mathscr{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{R})$ sa matrice dans une base orthonormale $\mathscr{B}$. La matrice $S = {}^tAA$ est symétrique (vérification facile) et alors :

$$\forall\, X,\, Y \in \mathcal{M}_{n,1}(\mathbb{R}),\ {}^tXSY = {}^tX{}^tAAY = {}^t(AX)AY = \langle u(x)\,|\,u(y)\rangle\ .$$

  En particulier,

$$\forall\, X \in \mathcal{M}_{n,1}(\mathbb{R}),\ {}^tXSX = \|u(x)\|^2\ .$$

Exemple

Soit $A = (a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ telle que $a_{ij} = \min(i,j)$ pour tout $(i,j) \in [\![1\,;n]\!]^2$.
Montrer que les valeurs propres de A sont strictement positives.

Nous proposons deux solutions, toutes deux instructives.

- Première méthode
  Pour tout $X = {}^t\begin{pmatrix}x_1 & \dots & x_n\end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$ on a :

$${}^tXAX = \sum_{1\leqslant i,j\leqslant n} a_{ij}x_ix_j = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\leqslant i<j\leqslant n} a_{ij}x_ix_j\ ,$$

  soit :

$$
\begin{aligned}
{}^tXAX &= \sum_{i=1}^{n} ix_i^2 + 2\sum_{1\leqslant i<j\leqslant n} ix_ix_j \\
&= (x_1+\cdots+x_n)^2 + \sum_{i=1}^{n}(i-1)x_i^2 + 2\sum_{1\leqslant i<j\leqslant n}(i-1)x_ix_j \\
&= (x_1+\cdots+x_n)^2 + \sum_{i=2}^{n}(i-1)x_i^2 + 2\sum_{2\leqslant i<j\leqslant n}(i-1)x_ix_j\,,
\end{aligned}
$$

et en réitérant le procédé on obtient :

$${}^tXAX = (x_1+\cdots+x_n)^2 + (x_2+\cdots+x_n)^2 + \cdots + (x_{n-1}+x_n)^2 + x_n^2\,.$$

On a donc, pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, ${}^tXAX \geqslant 0$, *et plus précisément, si* $X \neq 0$ *on aura* ${}^tXAX > 0$.
Si λ *est une valeur propre de* A, *il existe* $X \in \mathcal{M}_{n,1}(\mathbb{R})$, *non nul, tel que* $AX = \lambda X$; *on aura alors* ${}^tXAX = \lambda {}^tXX = \lambda \|X\|^2$ *et, puisque* $\|X\|^2 > 0$ *on en déduit* $\lambda > 0$.

- Seconde méthode

  La matrice A *s'écrit :* $A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \dots & n \end{pmatrix}$ *ce qui permet de remarquer que* $A = {}^tTT$ *où* T *est la matrice triangulaire supérieure*

$$T = \begin{pmatrix} 1 & \dots & \dots & 1 \\ & 1 & \dots & 1 \\ & 0 & \ddots & \vdots \\ & & & 1 \end{pmatrix}.$$

  On a donc, pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, ${}^tXAX = {}^tX{}^tTTX = \|TX\|^2$.
  Puisque T *est inversible, on a* $TX \neq 0$ *lorsque* $X \neq 0$ *donc* ${}^tXAX > 0$.
  On conclut alors comme ci-dessus.

SF6.5 Utiliser la stabilité de l'orthogonal

On rappelle ici une propriété qui peut être utile.

> Si u est un endomorphisme symétrique et si F est un sous-espace vectoriel stable par u, alors $F^\perp$ est aussi stable par u.

Cette propriété peut servir à faire des démonstrations par récurrence sur la dimension de l'espace vectoriel (elle est d'ailleurs utilisée dans la démonstration du *théorème spectral*).

Exemple

Soient p et q deux projecteurs orthogonaux d'un espace vectoriel euclidien E.

Montrer que E est somme directe orthogonale de sous-espaces vectoriels de dimensions 1 ou 2 stables par p et par q.

Procédons par récurrence sur $n = \dim E$.

- *Le résultat est immédiat si* $\dim E \leqslant 2$.

- *Supposons le résultat acquis pour tout espace vectoriel euclidien de dimension $\leqslant n$, et soit E de dimension $n+1$.*
  Si p et q sont des projecteurs orthogonaux de E, ce sont aussi des endomorphismes symétriques (cf. **[S5.9]***). Par conséquent, $p-q$ est aussi un endomorphisme symétrique, donc d'après le théorème spectral, il possède au moins une valeur propre réelle λ : il existe $x \neq 0$ tel que $(p-q)(x) = \lambda x$.*
  Notons $F = \mathrm{Vect}\big(x,p(x)\big)$ (F est une droite ou un plan). Alors :
  - *F est stable par p puisque $p[p(x)] = p(x)$.*
  - *F est stable par q : pour cela il suffit de montrer que $q(x)$ et $q[p(x)]$ sont dans F :*
    - *$q(x) = p(x) - \lambda x$ appartient à F ;*
    - *$q[p(x)] = q[q(x)+\lambda x] = q^2(x) + \lambda q(x) = (\lambda+1)q(x)$ appartient donc aussi à F.*

  p et q étant des endomorphismes symétriques, on en déduit que $F^\perp$ est aussi stable par p et par q. Il suffit donc d'appliquer l'hypothèse de récurrence aux endomorphismes induits par p et q sur $F^\perp$ pour conclure.

SF6.6 Utiliser le théorème spectral

Le théorème spectral est rappelé en **[S6.6]**, et il faut parfaitement le connaître, autant dans sa version algébrique que dans sa version matricielle. Il faut aussi bien connaître sa démonstration, dont les principes peuvent être utilisés dans d'autres situations.

✓ Le théorème spectral est incontournable : il faut y penser dans (presque) tous les problèmes mettant en jeu des endomorphismes symétriques.

Nous l'avons déjà utilisé plusieurs fois dans cet ouvrage. Nous en donnons ci-dessous d'autres exemples d'applications.

Exemple

Soit $S \in \mathcal{S}_n(\mathbb{R})$.
Montrer qu'il existe des réels $\lambda_1,\dots,\lambda_n$ et des matrices colonnes $U_1,\dots,U_n$ dans $\mathcal{M}_{n,1}(\mathbb{R})$ tels que :

$$\forall\, i \in [\![1\,;n]\!],\ {}^{\mathrm{t}}U_iU_i = 1 \ \text{ et } \ S = \sum_{i=1}^{n} \lambda_i U_i {}^{\mathrm{t}}U_i\,.$$

Comme d'habitude, nous confondons ici $\mathcal{M}_{n,1}(\mathbb{R})$ avec $\mathbb{R}^n$ muni de sa structure euclidienne canonique, et la matrice S avec l'endomorphisme qui lui est canoniquement associé.

D'après le théorème spectral, les valeurs propres $(\lambda_1,\dots,\lambda_n)$ de S sont toutes réelles et il existe une base orthonormale $(U_1,\dots,U_n)$ de $\mathbb{R}^n$ formée de vecteurs propres associés.

Les U_i étant unitaires, on a ${}^{\mathrm{t}}U_iU_i = \|U_i\|^2 = 1$ pour tout $i \in [\![1\,;n]\!]$. De plus,

$$\forall\, i \in [\![1\,;n]\!],\ SU_i = \lambda_i U_i$$

et

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \langle U_i \,|\, U_j\rangle = {}^{\mathrm{t}}U_iU_j = \delta_{ij}\,.$$

Si l'on note A la matrice $A = \sum_{i=1}^{n} \lambda_i U_i {}^t U_i$ on a donc :

$$\forall\, j \in [\![1\,;n]\!],\ AU_j = \sum_{i=1}^{n} \lambda_i U_i {}^t U_i U_j = \sum_{i=1}^{n} \lambda_i U_i \delta_{ij} = \lambda_j U_j = SU_j\,.$$

Ainsi les endomorphismes canoniquement associés à A et S coïncident sur une base donc sont égaux, d'où l'on déduit $S = A$ c'est-à-dire le résultat demandé.

Exemple

Soit $A \in \mathcal{M}_n(\mathbb{R})$, symétrique à valeurs propres positives, et $U \in \mathcal{M}_n(\mathbb{R})$, orthogonale.

Montrer que : $\mathrm{tr}(AU) \leqslant \mathrm{tr}(A)$.

A étant symétrique réelle, il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que :

$$^tPAP = P^{-1}AP = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)\,,$$

avec de plus $(\lambda_1, \ldots, \lambda_n) \in (\mathbb{R}_+)^n$.

On en déduit par propriété de la trace du produit de deux matrices :

$$\mathrm{tr}(AU) = \mathrm{tr}(PDP^{-1}U) = \mathrm{tr}(DP^{-1}UP)\,.$$

Notons $P^{-1}UP = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.

On a alors $DP^{-1}UP = \begin{pmatrix} \lambda_1 a_{11} & \cdots & \lambda_1 a_{1n} \\ \vdots & & \vdots \\ \lambda_n a_{1n} & \cdots & \lambda_n a_{nn} \end{pmatrix}$ *d'où* $\mathrm{tr}(AU) = \sum_{i=1}^{n} \lambda_i a_{ii}$.

De plus, $P^{-1}UP$ appartient à $\mathcal{O}_n(\mathbb{R})$ (car $\mathcal{O}_n(\mathbb{R})$ est un groupe pour la multiplication), donc ses vecteurs colonnes sont unitaires, c'est-à-dire $\sum_{i=1}^{n} a_{ij}^2 = 1$ pour tout j ; en particulier, tous les a_{ii} sont inférieurs à 1, d'où, les λ_i étant positifs, $\mathrm{tr}(AU) \leqslant \sum_{i=1}^{n} \lambda_i$, ce qui est le résultat voulu.

■ Comment montrer qu'un endomorphisme est une isométrie ?

SF6.7 Utiliser la définition

Pour montrer qu'un endomorphisme est une isométrie, il suffit de vérifier qu'il conserve la norme, ou le produit scalaire.

Exemple

$\mathcal{M}_n(\mathbb{R})$ est muni de son produit scalaire canonique. Soit $A \in \mathcal{M}_n(\mathbb{R})$, donnée.

À quelle condition l'application $f_A : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ M & \longmapsto & AM \end{cases}$ est-elle une isométrie ?

• *Il ne faut pas oublier de vérifier que f_A est bien un endomorphisme, ce qui découle directement des propriétés des lois dans $\mathcal{M}_n(\mathbb{R})$.*

• *f_A est une isométrie si et seulement si :*

$$\forall\, M, N \in \mathcal{M}_n(\mathbb{R}),\ \langle f_A(M) \,|\, f_A(N)\rangle = \langle M \,|\, N\rangle\ .$$

Or $\langle f_A(M) \,|\, f_A(N)\rangle = \mathrm{tr}({}^t(AM)AN) = \mathrm{tr}({}^tM{}^tAAN) = \langle M \,|\, {}^tAAN\rangle$ *donc :*

$$\begin{aligned} f_A \text{ isométrie} &\iff \forall\, M, N \in \mathcal{M}_n(\mathbb{R}),\ \left\langle M \,\middle|\, {}^tAAN\right\rangle = \langle M \,|\, N\rangle \\ &\iff \forall\, M, N \in \mathcal{M}_n(\mathbb{R}),\ \left\langle M \,\middle|\, {}^tAAN - N\right\rangle = 0 \\ &\iff \forall\, M, N \in \mathcal{M}_n(\mathbb{R}),\ {}^tAAN - N \perp M\,. \end{aligned}$$

Or dans un espace préhilbertien le seul vecteur orthogonal à tous les vecteurs est le vecteur nul donc la condition précédente équivaut à :

$$\forall\, N \in \mathcal{M}_n(\mathbb{R}),\ {}^tAAN = N$$

ce qui équivaut à : ${}^tAA = I_n$.

En conclusion : f_A est une isométrie $\iff$ $A \in \mathcal{O}_n(\mathbb{R})$.

Exemple

Soient $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_n)$ deux bases d'un espace vectoriel euclidien, telles que :

$$\forall\, (i,j) \in [\![1\,;n]\!]^2,\ \langle e_i \,|\, e_j\rangle = \langle f_i \,|\, f_j\rangle\,.$$

Montrer qu'il existe une isométrie $u \in \mathcal{O}(E)$ telle que $u(e_i) = f_i$ pour tout i.

On sait déjà qu'il existe un et un seul endomorphisme u de E qui transforme la base $(e_1, \ldots, e_n)$ en la base $(f_1, \ldots, f_n)$; il s'agit donc de montrer que u est une isométrie.

Pour cela, nous allons démontrer que u conserve la norme.

Soit $x \in E$. On le décompose dans la base $(e_1, \ldots, e_n)$: $x = \sum_{i=1}^{n} x_i e_i$, *d'où* $u(x) = \sum_{i=1}^{n} x_i u(e_i) = \sum_{i=1}^{n} x_i f_i$, *et on a alors :*

$$\|u(x)\|^2 = \left\langle \sum_{i=1}^{n} x_i f_i \,\middle|\, \sum_{j=1}^{n} x_j f_j \right\rangle = \sum_{1 \leqslant i,j \leqslant n} x_i x_j \,\langle f_i \,|\, f_j\rangle\ ,$$

et aussi :

$$\|x\|^2 = \left\langle \sum_{i=1}^{n} x_i e_i \,\middle|\, \sum_{j=1}^{n} x_j e_j \right\rangle = \sum_{1 \leqslant i,j \leqslant n} x_i x_j \,\langle e_i \,|\, e_j\rangle\ .$$

Compte tenu de l'hypothèse de l'énoncé, on a bien $\|u(x)\| = \|x\|$ pour tout $x \in E$, ce qui achève la démonstration.

SF6.8 Utiliser les matrices

Pour démontrer que u est une isométrie, on peut montrer que sa matrice dans une *base orthonormale* est une matrice orthogonale.

Pour montrer par le calcul qu'une matrice M est orthogonale, il est maladroit de faire le calcul de tous les termes de la matrice ${}^tMM$; il est plus rapide de démontrer que la famille de ses vecteurs colonnes (ou lignes) est orthonormale.

Exemple

Soit $A = -\frac{2}{3}\begin{pmatrix} -\frac{1}{2} & \frac{u}{v} & \frac{u}{w} \\ \frac{v}{u} & -\frac{1}{2} & \frac{v}{w} \\ \frac{w}{u} & \frac{w}{v} & -\frac{1}{2} \end{pmatrix}$.

Déterminer u, v, w tels que A soit la matrice d'une isométrie, puis caractériser cette isométrie.

En notant C_i les colonnes de la matrice A, on sait que A est orthogonale si et seulement si :

$$\forall\,(i,j) \in [\![1\,;3]\!]^2,\ \langle C_i \,|\, C_j\rangle = \delta_{ij}\,,$$

ce qui se traduit ici par les systèmes suivants (attention au facteur $-\frac{2}{3}$!) :

$$\begin{cases} \frac{1}{4} + \frac{v^2}{u^2} + \frac{w^2}{u^2} = \frac{9}{4} \\ \frac{u^2}{v^2} + \frac{1}{4} + \frac{w^2}{v^2} = \frac{9}{4} \\ \frac{u^2}{w^2} + \frac{v^2}{w^2} + \frac{1}{4} = \frac{9}{4} \end{cases} \quad et \quad \begin{cases} -\frac{u}{2v} - \frac{v}{2u} + \frac{w^2}{uv} = 0 \\ \frac{u^2}{vw} - \frac{v}{2w} - \frac{w}{2v} = 0 \\ -\frac{u}{2w} + \frac{v^2}{uw} - \frac{w}{2u} = 0 \end{cases},$$

qui se ramènent simplement à :

$$u^2 + v^2 = 2w^2 \quad et \quad v^2 + w^2 = 2u^2$$

c'est-à-dire : $u^2 = v^2 = w^2$. En posant $v = \varepsilon u$ et $w = \varepsilon' u$ avec ε, $\varepsilon' \in \{-1, +1\}$, on obtient donc que A est de la forme :

$$A = -\frac{2}{3}\begin{pmatrix} -\frac{1}{2} & \varepsilon & \varepsilon' \\ \varepsilon & -\frac{1}{2} & \varepsilon\varepsilon' \\ \varepsilon' & \varepsilon\varepsilon' & -\frac{1}{2} \end{pmatrix}.$$

On trouve alors facilement que l'ensemble des vecteurs invariants est le plan P d'équation $x + \varepsilon y + \varepsilon' z = 0$; d'après la classification des isométries rappelée en **[S6.4]**, *on peut dire sans plus de calculs que A est la matrice de la symétrie orthogonale par rapport à P.*

SF6.9 Utiliser l'image d'une base orthonormale

Pour démontrer que u est une isométrie, on peut montrer que l'image par u d'une base orthonormale est une base orthonormale.

Exemple

Quels sont les endomorphismes u de l'espace euclidien $\mathbb{R}^3$ qui conservent le produit vectoriel, c'est-à-dire tels que :

$$\forall\,(x,y) \in (\mathbb{R}^3)^2,\ u(x \wedge y) = u(x) \wedge u(y)\ ?$$

Soit f un endomorphisme de $\mathbb{R}^3$ qui conserve le produit vectoriel.
Soit $\mathscr{B} = (i,j,k)$ une base orthonormale directe de $\mathbb{R}^3$, et (I,J,K) son image par f. On a alors :

$$I \wedge J = f(i) \wedge f(j) = f(i \wedge j) = f(k) = K,$$

et de même :

$$J \wedge K = I \quad \text{et} \quad K \wedge I = J\,.$$

Distinguons deux cas.

- *Si l'un des vecteurs I, J ou K est nul, les relations ci-dessus impliquent $I = J = K = 0$, donc f est l'application nulle.*
- *Sinon la relation $K = I \wedge J$ montre que K est orthogonal à I et à J, et la relation $J \wedge K = I$ montre que I est orthogonal à J.*
  (I,J,K) est donc une famille orthogonale de vecteurs non nuls, donc une base, et c'est une base orthogonale directe compte tenu des relations ci-dessus.
  On en déduit : $\|K\| = \|I\|\,\|J\|$, $\|I\| = \|J\|\,\|K\|$ et $\|J\| = \|I\|\,\|K\|$ donc :

$$\|I\|\,\|J\|\,\|K\| = \left(\|I\|\,\|J\|\,\|K\|\right)^2,$$

  d'où $\|I\|\,\|J\|\,\|K\| = 1$ puis $\|I\|^2 = \|J\|^2 = \|K\|^2 = 1$.
  La famille (I,J,K) est donc une base orthonormale. Il en résulte que f est une isométrie. Puisque cette base est directe :

$$\det f = \det_{(i,j,k)}\big(f(i),f(j),f(k)\big) = \det_{(i,j,k)}(I,J,K) = 1\,,$$

  donc f est une isométrie positive, c'est-à-dire une rotation.

Réciproquement, il est clair que toute rotation conserve le produit vectoriel.

En conclusion, f conserve le produit vectoriel si et seulement si $f = 0$ ou f est une rotation.

■ Comment étudier une isométrie en dimension 2 ou 3 ?

SF6.10 Utiliser la matrice

Il s'agit ici d'utiliser tout ce qui a été dit en **[S6.3]** et **[S6.4]**.
Supposons donnée la matrice dans la base canonique d'un endomorphisme de $\mathbb{R}^2$ ou $\mathbb{R}^3$ que l'on suppose munis de leur structure euclidienne canonique.
Plusieurs approches sont possibles, mais un plan d'étude général peut être le suivant.

- Vérifier qu'il s'agit bien d'une matrice orthogonale.
- En dimension 2, le calcul du déterminant suffit à préciser la nature de l'isométrie ; on en cherche ensuite les éléments caractéristiques (angle pour une rotation, axe pour une réflexion).
- En dimension 3, le calcul du déterminant apporte un renseignement, mais peut ne pas être suffisant. Il faudra de toutes façons chercher les éléments propres : ce ne peut être que les sous-espaces propres associés aux valeurs propres éventuelles $+1$ et -1. Le fait de connaître la dimension du sous-espace des vecteurs invariants suffit à caractériser l'isométrie.

  Dans le cas d'une réflexion, c'est terminé puisque c'est nécessairement une réflexion par rapport au plan des vecteurs invariants ; dans le cas d'une rotation ou d'une antirotation, il reste à préciser l'angle.

La propriété suivante peut dans certains cas être utile et éviter des calculs :

Si A est une matrice orthogonale et symétrique, c'est la matrice d'une symétrie orthogonale.

En effet, si $A = {}^tA$ et ${}^tAA = I_n$ on a $A^2 = I_n$ donc A est la matrice d'une symétrie ; cette symétrie est orthogonale en vertu de la caractérisation rappelée en **[S5.9]**.

Exemple

Reconnaître l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est :

$$A = \frac{1}{7}\begin{pmatrix} -2 & 6 & -3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix}.$$

- *Notons C_1, C_2, C_3 les colonnes de A. Alors :*

$$\|C_1\|^2 = \|C_2\|^2 = \|C_3\|^2 = \frac{1}{49}(4+36+9) = 1$$

$$\langle C_1 \,|\, C_2\rangle = \langle C_2 \,|\, C_3\rangle = \langle C_1 \,|\, C_3\rangle = \frac{1}{49}(-12+18-6) = 0\,,$$

donc la matrice A est orthogonale.

- *La matrice A étant symétrique, c'est la matrice d'une symétrie orthogonale. Cherchons ses vecteurs invariants ; si $V = {}^t\begin{pmatrix} x & y & z \end{pmatrix}$, on a :*

$$AV = V \iff \begin{cases} -2x + 6y - 3z = 7x \\ 6x + 3y + 2z = 7y \\ -3x + 2y + 6z = 7z \end{cases} \iff 3x - 2y + z = 0\,.$$

Il s'agit de l'équation d'un plan ; d'après le cours, on peut donc directement conclure : A est la matrice de la réflexion par rapport au plan d'équation $3x - 2y + z = 0$.

✓ *Dans ces calculs, ne pas oublier le facteur « $\frac{1}{7}$ » devant la matrice !*

Exemple

Quel est l'endomorphisme de $\mathbb{R}^3$ de matrice $A = \frac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{pmatrix}$?

- *Comme dans l'exemple précédent, il est facile de vérifier que $A \in \mathcal{O}_3(\mathbb{R})$.*
- *Un calcul rapide (règle de Sarrus par exemple) montre que* $\det A = +1$. *Ainsi, A est la matrice d'une rotation.*
- *L'axe de cette rotation est l'ensemble des vecteurs invariants ; on résout, pour $V = {}^t\begin{pmatrix} x & y & z \end{pmatrix}$:*

$$AV = V \iff \begin{cases} \frac{1}{3}(2x + 2y - z) = x \\ \frac{1}{3}(-x + 2y + 2z) = y \\ \frac{1}{3}(2x - y + 2z) = z \end{cases},$$

ce qui équivaut à : $x = y = z$.
L'axe de la rotation est donc la droite D de base (u) avec $u = (1,1,1)$.

- *Si θ est l'angle de cette rotation, on a* $\operatorname{tr} A = 2\cos\theta + 1 = 2$ *soit* $\cos\theta = \frac{1}{2}$. *On a donc* $\theta = \pm\frac{\pi}{3}$ *modulo* 2π.
- *Le signe de θ dépend de l'orientation choisie sur D. Arbitrairement, orientons D par le vecteur* $e'_1 = \frac{u}{\|u\|}$. *Si v est un vecteur unitaire orthogonal à D, on a* $v \wedge f(v) = (\sin\theta)e'_1$ *(puisque l'angle* $\widehat{(v, f(v))}$ *est égal à θ), ce qui permet de trouver* $\sin\theta$ *(ce n'est pas la seule façon de trouver* $\sin\theta$*, mais c'est certainement la plus facile à retenir).*

  En choisissant par exemple : $v = \frac{1}{\sqrt{2}}(1,-1,0)$ *(le produit scalaire* $\langle u \,|\, v\rangle$ *est bien nul), on a :* $Av = \frac{1}{\sqrt{2}}(0,-1,1)$ *d'où :*

$$v \wedge Av = -\frac{1}{2}(1,1,1) = -\frac{\sqrt{3}}{2}e'_1 .$$

  Finalement, $\sin\theta = -\frac{\sqrt{3}}{2}$, *et l'angle de la rotation est* $-\frac{\pi}{3}$.
- *En conclusion, A est la matrice de la rotation d'axe dirigé et orienté par* $u = (1,1,1)$ *et d'angle* $\theta = -\frac{\pi}{3}$.

SF6.11 Utiliser les propriétés géométriques

Certains exercices exigent de combiner habilement les propriétés géométriques et algébriques des isométries.

Exemple

Soient f et g deux rotations de l'espace euclidien $\mathbb{R}^3$.

Déterminer une condition nécessaire et suffisante pour que $f \circ g = g \circ f$.

- *Si* $f = \mathrm{Id}$ *ou* $g = \mathrm{Id}$, *elles commutent et il n'y a rien à dire de plus.*
- *Sinon, f et g sont des rotations dont l'axe est une droite.*
  - *Si les axes de f et g sont les mêmes, elles commutent puisque leurs restrictions dans le plan orthogonal sont deux rotations planes, et des rotations planes commutent (car la composée de deux rotations planes d'angles θ_1 et θ_2 est une rotation plane d'angle $\theta_1 + \theta_2$).*
  - *Supposons maintenant que les axes de f et g sont distincts, et que* $f \circ g = g \circ f$. *Notons* $\mathrm{Vect}(u)$ *l'axe de f et* $\mathrm{Vect}(v)$ *l'axe de g, avec u et v unitaires et linéairement indépendants.*

    Puisque $f[g(u)] = g[f(u)] = g(u)$, *$g(u)$ est invariant par f ; il est donc colinéaire à u et, puisqu'il est aussi unitaire (car g est une isométrie) on a forcément* $g(u) = \pm u$. *Mais le cas* $g(u) = u$ *conduirait à u invariant par g donc colinéaire à v, ce qui est exclu. On a donc nécessairement* $g(u) = -u$; *cela implique que -1 est valeur propre de g ; puisque* $\det g = 1$, *les valeurs propres de g sont donc 1 (simple) et -1 (double) ce qui signifie que g est la symétrie orthogonale par rapport à la droite* $\mathbb{R}.v$ *(c'est aussi une rotation d'angle π, également appelée demi-tour).*
    De plus, les sous-espaces propres étant orthogonaux, on a nécessairement $u \perp v$.

De la même façon, on obtient que f est un demi-tour d'axe $\mathbb{R}.u$.
Ainsi, f et g sont deux demi-tours d'axes orthogonaux.

– *Il reste à examiner la réciproque : supposons que f et g sont deux demi-tours d'axes orthogonaux $\mathbb{R}.u$ et $\mathbb{R}.v$ respectivement (avec u et v unitaires). En posant $w = u \wedge v$, les matrices de f et g dans la base orthonormale (u,v,w) sont diagonales (avec un 1 et deux -1 sur la diagonale) donc commutent, ce qui prouve que $f \circ g = g \circ f$.*

• *Conclusion : f et g sont deux rotations qui commutent si et seulement si :*

$$\begin{cases} f = \mathrm{Id} \text{ ou } g = \mathrm{Id} \\ \text{ou } f \text{ et } g \text{ ont même axe} \\ \text{ou } f \text{ et } g \text{ sont deux demi-tours d'axes orthogonaux.} \end{cases}$$

À vous de jouer !

Exercice 6.1 **SF4.5** **SF6.6**

Soit $L \in \mathcal{M}_{1,n}(\mathbb{R})$ une matrice ligne non nulle.
Caractériser l'endomorphisme de $\mathbb{R}^n$ de matrice $A = {}^tLL$.

Exercice 6.2 **SF6.9**

Soit E un espace vectoriel euclidien de dimension $n \geqslant 1$, et $u \in \mathrm{GL}(E)$ tel que :

$$\forall\,(x,y) \in E,\ \langle x \,|\, y\rangle = 0 \implies \langle u(x) \,|\, u(y)\rangle = 0\,.$$

Montrer qu'il existe $\alpha > 0$ tel que : $\forall\, x \in E,\ \|u(x)\| = \alpha\,\|x\|$.
(*Indication :* on pourra considérer l'image d'une base orthonormale par u).

Exercice 6.3 **SF6.1** **SF6.6**

On munit $E = \mathbb{R}_n[X]$ du produit scalaire :

$$\forall\,(P,Q) \in \mathbb{R}_n[X]^2,\ \langle P \,|\, Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t\,.$$

Montrer que l'application : $u\colon P \mapsto (1 - X^2)P'' - 2XP'$ est un endomorphisme symétrique de E, et que ses valeurs propres sont négatives.

Exercice 6.4 **SF4.12** **SF6.2** **SF6.10**

Quelles sont les matrices A appartenant à $\mathcal{M}_2(\mathbb{R})$ telles que : ${}^tA = A^2$?

Exercice 6.5 **SF6.6**

Soit $A \in \mathcal{M}_n(\mathbb{R})$, nilpotente et telle que ${}^tAA = A{}^tA$. Montrer que $A = 0_n$.

Exercice 6.6 **SF2.24** **SF5.7** **SF6.7**

Soit E un espace euclidien et u un automorphisme orthogonal de E. On note $v = \mathrm{Id}_E - u$.

1) Montrer que : $\mathrm{Ker}\, v = (\mathrm{Im}\, v)^\perp$.

2) Soit p la projection orthogonale sur $\mathrm{Ker}\, v$. Montrer que :

$$\forall\, x \in E,\ \frac{1}{m}\sum_{k=0}^{m-1} u^k(x) \underset{m\to+\infty}{\longrightarrow} p(x)\,.$$

Exercice 6.7 **SF5.3** **SF6.8**

Soit $M = (m_{ij})_{1\leqslant i,j\leqslant n}$ une matrice orthogonale. Montrer que :

$$\left|\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\right| \leqslant n \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} |m_{ij}| \leqslant n^{3/2}\,.$$

Exercice 6.8 **SF6.4** **SF6.6**

On note $\mathcal{S}_n(\mathbb{R})$ l'espace vectoriel des matrices symétriques réelles.

Une matrice $A \in \mathcal{S}_n(\mathbb{R})$ est dite *positive* si : $\forall\, X \in \mathcal{M}_{n,1}(\mathbb{R})$, ${}^tXAX \geqslant 0$.

Elle est dite *définie positive* si : $\forall\, X \in \mathcal{M}_{n,1}(\mathbb{R})$, $X \neq 0 \Longrightarrow {}^tXAX > 0$.

L'ensemble des matrices symétriques réelles positives (respectivement définies positives) est noté $\mathcal{S}_n^+(\mathbb{R})$ (respectivement $\mathcal{S}_n^{++}(\mathbb{R})$).

1) Prouver que $A \in \mathcal{S}_n(\mathbb{R})$ est positive (respectivement définie positive) si et seulement si $\mathrm{Sp}\, A \subset \mathbb{R}_+$ (respectivement $\mathrm{Sp}\, A \subset \mathbb{R}_+^*$).

2) Prouver que $A \in \mathcal{S}_n(\mathbb{R})$ est positive si et seulement si il existe une matrice $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A = {}^tBB$. Caractériser de même les matrices symétriques définies positives.

3) Montrer que, pour toute matrice $A \in \mathcal{S}_n^+(\mathbb{R})$, il existe une et une seule matrice $R \in \mathcal{S}_n^+(\mathbb{R})$ telle que $R^2 = A$.

4) Démontrer que l'application $A \mapsto \exp A$ est une bijection de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.

Exercice 6.9 **SF6.11**

Soient E un espace euclidien, H et H' deux hyperplans de E et s (respectivement s') la réflexion par rapport à H (respectivement H').
Quand a t-on $s \circ s' = s' \circ s$?

Les endomorphismes d'un espace euclidien qui figurent au programme sont les endomorphismes symétriques et les isométries. Cependant deux autres types d'endomorphismes jouent un rôle important : ce sont les endomorphismes antisymétriques et les endomorphismes normaux.
La plupart de leurs propriétés peuvent se démontrer en s'inspirant des démonstrations faites en cours pour les endomorphismes symétriques ou les isométries. Ils font l'objet des deux exercices suivants, pour lesquels exceptionnellement aucune référence à un savoir-faire n'est indiquée.

Exercice 6.10

Soit u un endomorphisme d'un espace vectoriel euclidien tel que :

$$\forall\, x \in E,\ \langle u(x)\,|\,x\rangle = 0\,.$$

1) Montrer alors que :

$$\forall\, x,\, y \in E,\ \langle u(x)\,|\,y\rangle = -\,\langle x\,|\,u(y)\rangle\ .$$

u s'appelle un endomorphisme antisymétrique.

2) Montrer qu'un endomorphisme u de E est antisymétrique si et seulement si sa matrice dans une/toute base orthonormale de E est antisymétrique.

3) Si $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice antisymétrique, montrer que $\mathrm{Sp}_{\mathbb{C}}(A) \subset \mathrm{i}\mathbb{R}$.

4) **a)** Montrer que si u est un endomorphisme antisymétrique et si F est un sous-espace vectoriel de E stable par u, alors $F^{\perp}$ est stable par u.

b) Montrer que $A \in \mathcal{M}_n(\mathbb{R})$ est antisymétrique si et seulement si A est orthogonalement semblable à une matrice diagonale par blocs avec sur la diagonale des zéros et des blocs de la forme $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ où $a \in \mathbb{R}_+^*$.

Exercice 6.11

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite normale si elle commute avec sa transposée : $A\,{}^{\mathrm{t}}A = {}^{\mathrm{t}}A A$.
Un endomorphisme u d'un espace vectoriel euclidien est dit normal si sa matrice dans une base orthonormale est normale.

1) Si u est normal, montrer que sa matrice dans toute base orthonormale est normale.

2) Montrer que si u est un endomorphisme normal et si F est un sous-espace vectoriel de E stable par u, alors $F^{\perp}$ est stable par u.

3) Montrer que toute matrice normale d'ordre 2 est orthogonalement semblable à une matrice de la forme $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ avec $(\lambda,\mu) \in \mathbb{R}^2$ ou de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ avec $(a,b) \in \mathbb{R}^2$.

4) En utilisant le résultat de l'exercice **4.12** page 331, montrer qu'une matrice A est normale si et seulement si elle est orthogonalement semblable à une matrice diagonale par blocs avec sur la diagonale des réels et des blocs de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ où $(a,b) \in \mathbb{R}^2$.

Thème 7 - Espaces vectoriels normés

L'objectif du programme concernant les espaces vectoriels normés est de généraliser certaines notions vues en 1re année aux fonctions et aux suites vectorielles, et de se familiariser avec la notion de norme afin de préparer l'étude des suites et séries de fonctions.
Si beaucoup de raisonnements exposés ici ressemblent à ceux utilisés en analyse réelle, il faut se garder de croire qu'il suffit simplement de remplacer la valeur absolue par une norme. Les méthodes utilisées sont souvent plus subtiles, en particulier dans les espaces de dimension infinie où l'intuition que l'on pourrait avoir est souvent mise en défaut. Il faut donc faire particulièrement attention et, en dehors des exercices de simples vérifications de propriétés du cours, s'entraîner à faire des raisonnements rigoureux qui s'appuient sur une connaissance parfaite du cours.

■ Comment prouver qu'une application est une norme ?

Il s'agit d'un problème relativement simple puisqu'il s'agit de vérifier les propriétés de la définition. Cependant, on prendra soin de ne pas « bâcler » certaines vérifications, notamment celles qui mettent en jeu les notions de borne supérieure ou de maximum.

SF7.1 Revenir à la définition

Exemple

Montrer que N, définie sur $\mathbb{R}^2$ par :

$$\forall\,(x,y)\in\mathbb{R}^2,\ N(x,y)=\max\{|x|,|y|,|x-y|\}\,,$$

est une norme et représenter sa boule unité fermée.

- *On vérifie les propriétés d'une norme.*
- *N est clairement à valeurs dans $\mathbb{R}_+$.*
- *Soit $u\in\mathbb{R}^2$, $u=(x,y)$, et soit $\lambda\in\mathbb{R}$. Alors :*

$$N(\lambda u)=N(\lambda x,\lambda y)=\max\{|\lambda x|,|\lambda y|,|\lambda x-\lambda y|\}=|\lambda|N(u)\,.$$

- *Soit $u\in\mathbb{R}^2$, $u=(x,y)$, alors :*

$$N(u)=0\iff\max\{|x|,|y|,|x-y|\}=0\iff x=y=0\iff u=0_{\mathbb{R}^2}\,.$$

- *Soit $(u,v)\in(\mathbb{R}^2)^2$, $u=(x,y)$, $v=(x',y')$. Alors :*

$$N(u+v)=N(x+x',y+y')=\max\{|x+x'|,|y+y'|,|x+x'-y-y'|\}\,.$$

$$\textit{Or :}\ \begin{cases} |x+x'| & \leqslant & |x|+|x'| & \leqslant & N(u)+N(v)\,,\\ |y+y'| & \leqslant & |y|+|y'| & \leqslant & N(u)+N(v)\,,\\ |x+x'-y-y'| & \leqslant & |x-y|+|x'-y'| & \leqslant & N(u)+N(v)\,, \end{cases}$$

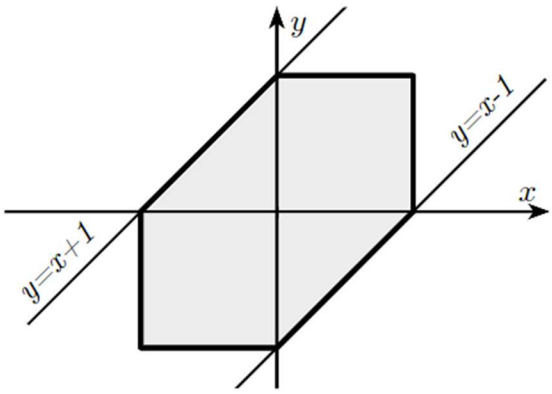
et donc $N(u+v)\leqslant N(u)+N(v)$.

En conclusion : N est une norme sur $\mathbb{R}^2$.

• *La boule unité fermée B_f est caractérisée par :*

$$(x,y) \in B_f \iff N(x,y) \leqslant 1$$
$$\iff \begin{cases} -1 \leqslant x \leqslant 1 \\ -1 \leqslant y \leqslant 1 \\ -1 \leqslant x - y \leqslant 1 \end{cases}$$

d'où la figure ci-contre.



SF7.2 Utiliser l'algèbre bilinéaire

Cette méthode consiste à repérer le cas d'une norme qui est associée à un produit scalaire. Les calculs (principalement l'inégalité triangulaire) en seront simplifiés.

On peut penser à l'appliquer lorsque, dans l'expression de la norme, figurent des carrés et des radicaux.

Exemple

Soit $E = \mathbb{R}_n[X]$ et $a_0, \ldots, a_n$ $n+1$ réels distincts.

Pour tout $P \in E$ on pose $N(P) = \sqrt{\sum_{k=0}^{n} P(a_k)^2}$.

Montrer que N est une norme sur E.

Il est ici naturel de considérer l'application :

$$\varphi : (P,Q) \in E^2 \longmapsto \sum_{k=0}^{n} P(a_k)Q(a_k)\,.$$

Il est alors facile de vérifier que φ est une forme bilinéaire symétrique positive ; de plus, φ est définie car :

$$\varphi(P,P) = 0 \implies \sum_{k=0}^{n} P(a_k)^2 = 0 \implies \forall\, k \in [\![1\,;n]\!],\ P(a_k) = 0 \implies P = 0\,,$$

car si un polynôme de $\mathbb{R}_n[X]$ s'annule en $n+1$ points distincts, il est nul. Donc φ est un produit scalaire sur E,

On a alors, pour tout $P \in E$, $N(P) = \sqrt{\varphi(P,P)}$; ainsi N est la norme euclidienne associée au produit scalaire φ, donc c'est une norme.

■ Comment comparer deux normes ?

SF7.3 Revenir à la définition

Pour comparer deux normes N_1 et N_2 sur un espace vectoriel normé E, on cherche s'il existe des réels α et β strictement positifs tels que $N_1(x) \leqslant \alpha N_2(x)$ et $N_2(x) \leqslant \beta N_1(x)$ pour tout $x \in E$; certaines inégalités ne sont d'ailleurs pas toujours faciles à trouver.

Exemple

Soit $E = \mathscr{C}^1\big([0\,;1],\mathbb{R}\big)$. Pour toute $f \in E$ on pose :

$$N_1(f) = |f(0)| + \int_0^1 |f'(t)|\,\mathrm{d}t \quad \text{et} \quad N_2(f) = |f(1)| + \int_0^1 |f'(t)|\,\mathrm{d}t\,.$$

Montrer qu'il s'agit de normes sur E et les comparer.

- *Le fait que N_1 et N_2 sont des normes sur E se vérifie facilement, et est laissé à titre d'exercice.*
- *Pour relier les deux normes il faut penser à la relation :*

$$\forall\, f \in E,\ f(1) = f(0) + \int_0^1 f'(t)\,\mathrm{d}t\,.$$

On en déduit :

$$\forall\, f \in E,\ |f(1)| \leqslant |f(0)| + \left|\int_0^1 f'(t)\,\mathrm{d}t\right| \leqslant |f(0)| + \int_0^1 |f'(t)|\,\mathrm{d}t = N_1(f)$$

donc :

$$\forall\, f \in E,\ N_2(f) = |f(1)| + \int_0^1 |f'(t)|\,\mathrm{d}t \leqslant N_1(f) + N_1(f) = 2N_1(f)\,.$$

À partir de la relation $f(0) = f(1) - \displaystyle\int_0^1 f'(t)\,\mathrm{d}t$, on obtient exactement de la même façon :

$$\forall\, f \in E,\ N_1(f) \leqslant 2N_2(f)\,.$$

En conclusion : $\dfrac{1}{2}N_2 \leqslant N_1 \leqslant 2N_2$, et les deux normes sont équivalentes.

SF7.4 Utiliser des suites

Dire que deux normes N_1 et N_2 dans un espace vectoriel normé E sont équivalentes signifie que les rapports $\dfrac{N_1}{N_2}$ et $\dfrac{N_2}{N_1}$ sont bornés sur $E \setminus \{0\}$.

Pour montrer qu'elles ne sont pas équivalentes, on peut donc essayer de trouver des suites $(x_n) \in E^{\mathbb{N}}$ telles que $\displaystyle\lim_{n\to+\infty} \frac{N_1(x_n)}{N_2(x_n)} = +\infty$ ou $\displaystyle\lim_{n\to+\infty} \frac{N_2(x_n)}{N_1(x_n)} = +\infty$.

On peut aussi essayer de trouver des suites qui n'ont pas les mêmes propriétés pour les deux normes, par exemple qui sont bornées au sens de l'une mais pas de l'autre, ou qui convergent vers des limites différentes selon la norme considérée.

Exemple

On note $E = \big\{f \in \mathscr{C}^2\big([0\,;1],\mathbb{R}\big) \;\big|\; f(0) = f(1) = 0\big\}$. On définit l'application N sur E par $N(f) = \left(\displaystyle\int_0^1 f''(t)^2\,\mathrm{d}t\right)^{1/2}$.

1) Montrer que N est une norme.

2) Comparer $\|\cdot\|_\infty$ et N.

1) *Notons déjà que E est bien un $\mathbb{R}$-espace vectoriel (comme sous-espace vectoriel de $\mathscr{C}^2\big([0\,;1],\mathbb{R}\big)$).*

Si $N(f) = 0$ alors $f'' = 0$ donc $f' = cste$, et puisque $f(0) = f(1) = 0$, $f = 0$. Les autres propriétés d'une norme sont faciles à vérifier, en remarquant que $N(f) = \|f''\|_2$, et en utilisant le fait que $\|\cdot\|_2$ est une norme d'après le cours.

2) • *Soit $f \in E$. On a alors, pour tout $x \in [0\,;1]$, en utilisant l'inégalité de Cauchy-Schwarz :*

$$|f'(x) - f'(0)| = \left|\int_0^x f''(t)\cdot 1\,\mathrm{d}t\right| \leqslant \left(\int_0^x f''(t)^2\,\mathrm{d}t\right)^{1/2}\cdot\sqrt{x} \leqslant N(f)\,.$$

Or la fonction f' s'annule en au moins un point $c \in [0\,;1]$ d'après le théorème de Rolle, donc l'inégalité précédente appliquée avec $x = c$ donne :

$$|f'(0)| \leqslant N(f)\,.$$

La même inégalité implique alors, d'après l'inégalité triangulaire :

$$\forall\, x \in [0\,;1],\ |f'(x)| \leqslant 2N(f)\,,$$

et enfin l'inégalité des accroissement finis permet de conclure :

$$\|f\|_\infty \leqslant 2N(f)\,.$$

• *Considérons, pour $n \in \mathbb{N}^*$, $f_n\colon x \mapsto \sin(n\pi x)$. Alors $\|f_n\|_\infty = 1$ mais $N(f_n) = \dfrac{1}{\sqrt{2}}\,n^2\pi^2$ tend vers $+\infty$ avec n. Ces deux normes ne sont donc pas équivalentes.*

■ Comment étudier une suite dans un espace vectoriel normé ?

SF7.5 Revenir à la définition

On peut revenir à la définition de la convergence d'une suite u vers ℓ à l'aide des ε pour résoudre certains exercices théoriques : c'est ce que nous ferons par exemple en **[SF8.14]** dans le cas de suites réelles. Cette méthode peut être assez délicate, mais elle est importante. C'est pourquoi nous en donnons plusieurs exemples.

Exemple

Montrer que si une suite $(u_n)_{n\in\mathbb{N}}$ d'un compact K possède une unique valeur d'adhérence, elle est convergente.

Soit donc u une suite d'éléments de K n'ayant qu'une seule valeur d'adhérence ℓ, et montrons que u converge vers ℓ.
Pour cela, nous allons raisonner par l'absurde, en supposant que (u_n) ne converge pas vers ℓ ; par négation de la définition de la limite, cela signifie :

$$\exists\,\varepsilon > 0,\ \forall\, N \in \mathbb{N},\ \exists\, n \geqslant N \quad \textit{tel que}\ \|u_n - \ell\| > \varepsilon\,.$$

En appliquant cette proposition avec $N = 0$, on obtient l'existence d'un entier n_0 tel que $\|u_{n_0} - \ell\| > \varepsilon$; puis en l'appliquant de nouveau avec $N = n_0 + 1$, on obtient l'existence d'un entier $n_1 > n_0$ tel que $\|u_{n_1} - \ell\| > \varepsilon$ etc...

On peut donc construire une suite extraite $(u_{\varphi(n)})$ telle que :

$$\forall\, n \in \mathbb{N},\ \left\|u_{\varphi(n)} - \ell\right\| > \varepsilon\,.$$

Mais cette suite étant formée d'éléments du compact K possède nécessairement une valeur d'adhérence, qui ne peut être que ℓ d'après l'hypothèse. Cela est absurde car tous les termes de cette suite vérifient $\left\|u_{\varphi(n)} - \ell\right\| > \varepsilon$.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ une suite réelle telle que $\lim\limits_{n\to+\infty}(u_{n+1}-u_n)=0$.

Montrer que l'ensemble des valeurs d'adhérence de u est un intervalle de $\mathbb{R}$.

Si l'ensemble des valeurs d'adhérence de u est vide ou réduit à un point, le résultat est immédiat.

Sinon, nous allons démontrer que cet ensemble est une partie convexe (on sait que les parties convexes de $\mathbb{R}$ sont exactement les intervalles).
Soient donc $a<b$ deux valeurs d'adhérence de la suite u, et $c\in\,]a\,;b[$; il s'agit de montrer que c est encore une valeur d'adhérence, ce qui équivaut à :

$$\forall\,\varepsilon>0,\ \forall\, N\in\mathbb{N},\ \exists\, n\geqslant N \ \text{ tel que } |u_n-c|\leqslant\varepsilon$$

(en effet, cette propriété permet facilement, comme dans l'exemple précédent, de construire une suite extraite $(u_{\varphi(n)})$ qui converge vers c).
On raisonne alors par l'absurde, en supposant :

$$\exists\,\varepsilon>0,\ \exists\, N\in\mathbb{N},\ \forall\, n\geqslant N,\ u_n\notin[c-\varepsilon\,;c+\varepsilon]\,.$$

a et b étant des valeurs d'adhérence, on a donc $a\leqslant c-\varepsilon$ et $b\geqslant c+\varepsilon$.
Puisque $\lim\limits_{n\to+\infty}(u_{n+1}-u_n)=0$ on a aussi, par la définition de la limite :

$$\exists\, N_1\in\mathbb{N},\ n\geqslant N_1 \implies |u_{n+1}-u_n|<\varepsilon\,.$$

Puisque a est une valeur d'adhérence, il existe $n_0\geqslant \max(N,N_1)$ tel que $|u_{n_0}-a|<\varepsilon$. On aura alors $u_{n_0}<a+\varepsilon\leqslant c$ et puisque $u_{n_0}\notin[c-\varepsilon\,;c+\varepsilon]$, on a $u_{n_0}<c-\varepsilon$. Ensuite, $u_{n_0+1}<u_{n_0}+\varepsilon\leqslant c$ et pour la même raison, on en déduit $u_{n_0+1}<c-\varepsilon$.
Par récurrence, on a donc $u_n<c-\varepsilon$ pour tout $n\geqslant n_0$, ce qui contredit le fait que b est aussi une valeur d'adhérence de la suite, et achève la démonstration.

✓ Notez dans les deux exemples précédents l'utilisation de démonstrations par l'absurde, donc l'importance de savoir écrire la négation d'une proposition quantifiée.

Plus simplement, on peut aussi dans certains cas utiliser la caractérisation :

$$\lim_{n\to+\infty} u_n=\ell \iff \lim_{n\to+\infty}\|u_n-\ell\|=0$$

qui permet de ramener le problème de la convergence de la suite u vers ℓ à celui d'une suite réelle.

Exemple

Soit $E=\mathscr{C}\big([0\,;1],\mathbb{R}\big)$ et pour tout $n\in\mathbb{N}^*$, soit $f_n : t\in[0\,;1]\mapsto t^n$.

Montrer que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers la fonction nulle au sens de la norme $\|\cdot\|_1$ mais pas au sens de la norme $\|\cdot\|_\infty$.

Notons f la fonction nulle sur $[0\,;1]$.
Par définition, dire que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f au sens de la norme $\|\cdot\|_1$ signifie :

$$\lim_{n\to+\infty} \|f_n - f\|_1 = 0\,.$$

On calcule donc :

$$\|f_n - f\|_1 = \int_0^1 |t^n - 0|\,\mathrm{d}t = \int_0^1 t^n = \frac{1}{n+1}\,,$$

ce qui donne le résultat voulu.
Cependant, $\|f_n - f\|_\infty = \sup_{t\in[0;1]} t^n = 1$, donc $(f_n)_{n\in\mathbb{N}}$ ne converge pas vers f pour cette norme.

✓ *Cet exemple illustre le fait que les deux normes considérées sur $\mathscr{C}\big([0\,;1],\mathbb{R}\big)$, qui n'est pas de dimension finie, ne sont pas équivalentes.*

SF7.6 Utiliser les suites coordonnées

Si une suite est à valeurs dans un espace vectoriel de dimension finie, elle converge si et seulement si ses suites coordonnées dans une base quelconque convergent (et les coordonnées de la limite sont les limites des coordonnées). Ce critère permet donc de se ramer à des suites à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

Exemple

Soit $a \in \mathbb{R}$. Déterminer : $\displaystyle\lim_{n\to+\infty} \begin{pmatrix} 1 & -\frac{a}{n} \\ \frac{a}{n} & 1 \end{pmatrix}^n$.

Notons $A = \begin{pmatrix} 1 & -\frac{a}{n} \\ \frac{a}{n} & 1 \end{pmatrix}$. Nous allons tout simplement calculer A^n.
On remarque que la matrice A est « presque » orthogonale, puisque ses colonnes sont orthogonales et de mêmes normes ; plus précisément, on peut écrire :

$$A = \sqrt{1+\frac{a^2}{n^2}}\,B$$

où B est une matrice orthogonale (ses colonnes forment une base orthonormale de $\mathbb{R}^2$), de déterminant positif. B est donc une matrice de rotation, et l'on peut écrire :

$$B = \begin{pmatrix} \cos\theta_n & -\sin\theta_n \\ \sin\theta_n & \cos\theta_n \end{pmatrix} \text{ avec } \cos\theta_n = \frac{1}{\sqrt{1+\frac{a^2}{n^2}}} \text{ et } \sin\theta_n = \frac{\frac{a}{n}}{\sqrt{1+\frac{a^2}{n^2}}}\cdot$$

Or il est facile de calculer les puissances d'une matrice de rotation, puisque le produit de deux matrices de rotations d'angle respectifs θ et θ' est une matrice de rotation d'angle $\theta+\theta'$.
On aura donc :

$$B^n = \begin{pmatrix} \cos n\theta_n & -\sin n\theta_n \\ \sin n\theta_n & \cos n\theta_n \end{pmatrix} \quad \text{puis} \quad A^n = \left(1+\frac{a^2}{n^2}\right)^{\frac{n}{2}} B^n\,.$$

Or $\cos\theta_n > 0$ *donc on peut prendre* $\theta_n \in \left]-\frac{\pi}{2}\,;\frac{\pi}{2}\right[$ *de sorte que* $\theta_n = \operatorname{Arctan}\frac{a}{n}$. *Puisque* $\operatorname{Arctan} x \underset{x\to 0}{\sim} x$ *on en déduit* $\lim\limits_{n\to+\infty} n\theta_n = a$.

On en conclut que la suite de matrices (B^n) *converge vers la matrice* $\begin{pmatrix}\cos a & -\sin a\\ \sin a & \cos a\end{pmatrix}$ *(en faisant la limite pour chaque élément de la matrice).*

Enfin,

$$\lim_{n\to+\infty}\left(1+\frac{a^2}{n^2}\right)^{\frac{n}{2}} = \lim_{n\to+\infty} \mathrm{e}^{\frac{n}{2}\ln\left(1+\frac{a^2}{n^2}\right)} = 1$$

(puisque $\frac{n}{2}\ln\left(1+\frac{a^2}{n^2}\right) \underset{n\to+\infty}{\sim} \frac{a^2}{2n}$*).*

En conclusion : $\lim\limits_{n\to+\infty} A^n = \begin{pmatrix}\cos a & -\sin a\\ \sin a & \cos a\end{pmatrix}$.

SF7.7 Utiliser la caractérisation séquentielle de la continuité

Lors de l'étude d'une suite, on est souvent amené à utiliser le résultat suivant.

> Si une suite $(u_n)_{n\in\mathbb{N}}$ converge vers ℓ et si f est une application continue en ℓ, alors la suite $(f(u_n))_{n\in\mathbb{N}}$ converge vers $f(\ell)$.

Vous avez certainement l'habitude d'utiliser cette propriété lors de l'étude d'une suite réelle ; donnons un exemple d'utilisation pour l'étude de suites de matrices.

Exemple

1) Montrer que toute matrice de $\mathcal{M}_p(\mathbb{C})$ est limite d'une suite de matrices ayant toutes p valeurs propres distinctes.

2) En considérant la matrice $A = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$, montrer que ce résultat n'est pas vrai dans $\mathcal{M}_2(\mathbb{R})$.

1) *Soit* $A \in \mathcal{M}_p(\mathbb{C})$. *On sait que* A *est trigonalisable : il existe* $P \in \mathrm{GL}_n(\mathbb{C})$ *et* $T \in \mathcal{T}_n^+(\mathbb{C})$ *telles que* $A = PTP^{-1}$.
Notons $\lambda_1,\dots,\lambda_n$ *les termes diagonaux de* T *(c'est-à-dire les valeurs propres de* A*), et, pour* $n \in \mathbb{N}^*$, $T_n = T + \operatorname{diag}\left(\frac{1}{n},\frac{2}{n},\dots,\frac{p}{n}\right)$.

– *si pour* $i \neq j \in [\![1\,;p]\!]$ *on a* $\lambda_i = \lambda_j$, *alors* $\lambda_i + \frac{i}{n} \neq \lambda_j + \frac{j}{n}$;

– *si* $\lambda_i \neq \lambda_j$, *alors l'égalité* $\lambda_i + \frac{i}{n} = \lambda_j + \frac{j}{n}$ *ne peut avoir lieu que si* $n = \frac{i-j}{\lambda_j - \lambda_i}$, *donc pour au plus une valeur de* n.

À partir d'un certain rang, les matrices T_n *auront donc tous leurs termes diagonaux distincts.*
Il est clair que $\lim\limits_{n\to+\infty} T_n = T$ *(en faisant les limites de chaque coefficient des matrices). Puisque l'application* $M \mapsto PMP^{-1}$ *est linéaire sur un espace vectoriel de dimension finie, elle est continue donc :*

$$\lim_{n\to+\infty} PT_nP^{-1} = PTP^{-1} = A\,.$$

La suite des matrices PT_nP^{-1} *(à partir d'un certain rang) répond donc bien à la question posée.*

✓ *Une conséquence du résultat que nous venons de démontrer est que l'ensemble des matrices diagonalisables est dense dans* $\mathcal{M}_p(\mathbb{C})$.

2) *Supposons par l'absurde que la matrice* A *de l'énoncé soit limite, dans* $\mathcal{M}_2(\mathbb{R})$, *d'une suite de matrices* (A_n) *ayant toutes 2 valeurs propres distinctes. Cela signifie que, pour tout* n, *le polynôme caractéristique* χ_{A_n} *de* A_n *possède deux racines réelles distinctes, donc est à discriminant* $\Delta(A_n)$ *positif. Or pour une matrice* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, *le discriminant* $\Delta(M)$ *de son polynôme caractéristique vaut* $(a+d)^2 - 4bc$, *donc est une fonction continue de* M.
On devrait donc avoir : $\Delta(A) = \lim\limits_{n\to+\infty} \Delta(A_n) \geqslant 0$ *; mais* $\Delta(A) < 0$ *car le polynôme caractéristique de* A *vaut* X^2+1, *d'où la contradiction cherchée.*

■ Comment montrer qu'une partie est ouverte ?

SF7.8 Utiliser la définition

Pour montrer qu'une partie A est un ouvert, on peut montrer, en général à l'aide d'inégalités sur les distances, que tout point de A est intérieur à A.

Exemple

Montrer que l'ensemble : $A = \{(x_1,\dots,x_n) \in \mathbb{R}^n \mid x_i \neq x_j \text{ pour tous } i \neq j\}$ est un ouvert de $\mathbb{R}^n$.

Toutes les normes sur $\mathbb{R}^n$ *étant équivalentes, on utilisera ici la norme* $\|\cdot\|_\infty$.
Soit $a = (a_1,\dots,a_n) \in A$. *Posons* $r = \frac{1}{2}\min\limits_{i\neq j} |a_i - a_j|$. *Puisque* $a \in A$, r *est strictement positif.*
Pour tout $x = (x_1,\dots,x_n) \in \mathcal{B}(a,r)$ *et tout* $i \neq j$, *on a :*

$$\begin{aligned} |x_i - x_j| &= |x_i - a_i + a_i - a_j + a_j - x_j| \\ &\geqslant \Big|\, |a_i - a_j| - |(x_i - a_i) - (x_j - a_j)| \,\Big| . \end{aligned}$$

Or $|a_i - a_j| \geqslant 2r$ *et*

$$|(x_i - a_i) - (x_j - a_j)| \leqslant |x_i - a_i| + |x_j - a_j| \leqslant 2\,\|x - a\|_\infty < 2r\,,$$

donc $|x_i - x_j| > 0$ *et* $x \in A$. *Ainsi* $\mathcal{B}(a,r) \subset A$ *: pour tout* $a \in A$ *il existe une boule ouverte de centre* a *qui est incluse dans* A, A *est un ouvert.*

SF7.9 Utiliser certaines propriétés des ouverts

Il s'agit d'utiliser les propriétés : une réunion d'une famille quelconque d'ouverts est un ouvert, et une intersection d'une famille finie d'ouverts est un ouvert.

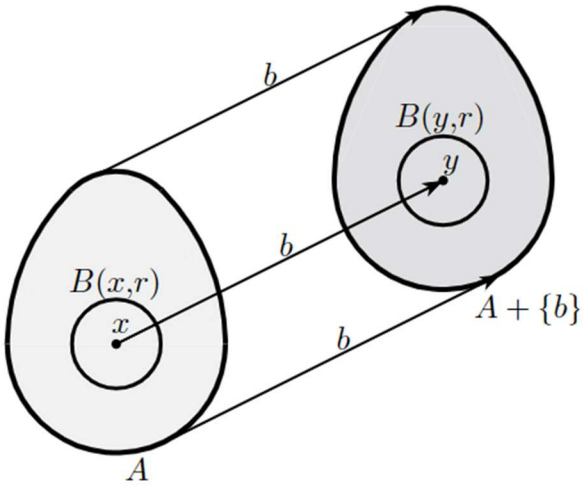
Exemple

Soit A un ouvert non vide de E, et B une partie non vide de E.

Montrer que $A + B = \{a + b \mid (a,b) \in A \times B\}$ est un ouvert.

- *Soit A un ouvert, et $b \in E$. Montrons déjà que $A+\{b\} = \{a+b \mid a \in A\}$, image de A par la translation de vecteur b, est un ouvert.*

  Soit $y \in A+\{b\}$. Il existe $x \in A$ tel que $y = x+b$.
  A étant ouvert, il existe $r > 0$ tel que $B(x,r) \subset A$.
  Il est alors facile de vérifier que $B(x+b,r) = B(x,r)+\{b\}$, donc $B(y,r) \subset A+\{b\}$, et $A+\{b\}$ est un ouvert.



- *Dans le cas général, $A+B = \bigcup_{b\in B}\Big(A+\{b\}\Big)$ est la réunion d'ouverts, donc c'est un ouvert.*

SF7.10 Utiliser une application continue bien choisie

Il s'agit certainement de la méthode la plus puissante et de la plus élégante : l'image réciproque d'une partie ouverte (respectivement fermée) par une application continue est encore une partie ouverte (respectivement fermée).

Exemple

Montrer que $\mathrm{GL}_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$.

$\mathrm{GL}_n(\mathbb{K})$ est l'ensemble des matrices de déterminant non nul.
C'est donc l'image réciproque de $\mathbb{R}^*$ par l'application det : $\mathcal{M}_n(\mathbb{K}) \to \mathbb{R}$.
Or $\mathbb{R}^*$ est un ouvert (car par exemple son complémentaire $\{0\}$ est une boule fermée donc un fermé), et l'application det *est continue.*
Par théorème, $\mathrm{GL}_n(\mathbb{K})$ image réciproque d'un ouvert par une application continue, est un ouvert.

✓ *Cela signifie intuitivement que dans un voisinage suffisamment petit d'une matrice inversible il n'y a que des matrices inversibles, ou encore que si l'on modifie suffisamment peu les coefficients d'une matrice inversible, elle reste inversible.*

■ Comment montrer qu'une partie est fermée ?

SF7.11 Utiliser la définition

Un fermé est, par définition, le complémentaire d'un ouvert. Compte tenu de la définition d'un ouvert, on a donc :

$$F \text{ fermé} \iff \forall x \notin F,\ \exists r > 0 \text{ tel que } \mathcal{B}(x,r) \cap F = \emptyset\,.$$

Exemple

Montrer qu'une boule fermée est un fermé.

Soit une boule fermée $\mathcal{B}_f(a,r)$. Si $x \notin \mathcal{B}_f(a,r)$ alors $\|x-a\| > r$. Posons alors $r' = \dfrac{\|x-a\|-r}{2}$, et montrons que $\mathcal{B}(x,r') \cap \mathcal{B}_f(a,r) = \emptyset$, ce qui démontrera le résultat voulu (nous vous invitons à faire une figure pour illustrer cette démonstration).
En effet, par l'absurde, s'il existait $y \in \mathcal{B}(x,r') \cap \mathcal{B}_f(a,r)$ on aurait, à l'aide de l'inégalité triangulaire :

$$\|x-a\| = \|(x-y)+(y-a)\| \leqslant \underbrace{\|x-y\|}_{<r'} + \underbrace{\|y-a\|}_{\leqslant r}$$

$$< r + r' = \frac{\|x-a\|+r}{2} < \|x-a\| \ ,$$

d'où la contradiction.

✓ Cette méthode est en fait très peu utilisée et on lui préfèrera les deux méthodes suivantes. Par exemple ici, il était plus facile de répondre à la question en disant simplement que $\mathcal{B}_f(a,r)$ est l'image réciproque du fermé $[0\,;r]$ par l'application continue (car 1-lipschitzienne) $x \mapsto \|x-a\|$, ou en remarquant que, si (x_n) est une suite de vecteurs qui converge vers x et tels que $\|x_n - a\| \leqslant r$, on a encore $\|x-a\| \leqslant r$.

SF7.12 Utiliser la caractérisation séquentielle

La caractérisation séquentielle d'un fermé est certainement la méthode la plus intéressante pour caractériser un fermé dans un espace vectoriel normé. Cette méthode, dont il n'existe pas d'équivalent pour les ouverts contrairement aux autres méthodes étudiées ici, fait qu'il est souvent plus facile de montrer qu'une partie est fermée que de montrer qu'une partie est ouverte. Son avantage est de pouvoir utiliser tous les outils concernant les suites.

Exemple

Soit F un sous-espace vectoriel d'un espace vectoriel normé E de dimension finie. Montrer que F est un fermé.

Soit $(e_1,\dots,e_p)$ une base de F, que l'on complète en une base $\mathscr{B} = (e_1,\dots,e_n)$ de E.
Considérons une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de F, qui converge vers $\ell \in E$. En écrivant les coordonnées de ces vecteurs dans la base $\mathscr{B}$ on a :

$$\forall\, n \in \mathbb{N},\ x_n = \sum_{i=1}^{n} x_n^{(i)} e_i \quad \text{et} \quad \ell = \sum_{i=1}^{n} \ell_i e_i \ ,$$

avec, pour tout $i \in [\![1\,;n]\!]$: $\lim\limits_{n\to+\infty} x_n^{(i)} = \ell_i$.

Mais pour tout n les $x_n^{(i)}$ sont nuls pour $i \geqslant p+1$ puisque $x_n \in F$. On aura donc aussi $\ell_i = 0$ pour $i \geqslant p+1$, donc $\ell \in F$.

Ainsi, la limite de toute suite convergente d'éléments de F est dans F : F est fermé.

SF7.13 Utiliser une application continue bien choisie

Comme pour les ouverts, il s'agit d'utiliser l'image réciproque d'une partie par une application continue.

Exemple

Soit H un hyperplan d'un espace vectoriel normé E.

1) Montrer que si E est de dimension finie, H est un fermé.

2) Dans le cas général, montrer que H est soit fermé, soit dense dans E.

1) *H étant un hyperplan, il existe une forme linéaire φ non nulle telle que $H = \operatorname{Ker}\varphi$, soit $H = \varphi^{-1}(\{0\})$. φ étant une application linéaire définie sur un espace vectoriel de dimension finie, elle est continue. H est donc l'image réciproque du fermé $\{0\}$ de $\mathbb{R}$ par φ donc est un fermé.*

2) *Dans le cas général, supposons que H n'est pas fermé, et montrons que H est dense dans E, c'est-à-dire que tout élément de E est limite d'une suite d'éléments de H.*
H n'étant pas fermé, d'après la caractérisation séquentielle (ou plutôt sa négation), on peut trouver une suite convergente $(x_n)_{n\in\mathbb{N}}$ d'éléments de H et un élément $a \notin H$ tels que $\lim\limits_{n\to+\infty} x_n = a$.
Puisque $a \notin H$, on sait que $E = H \oplus \mathbb{K}.a$ (voir **[S2.20]***).*
Si b est un vecteur quelconque de E, on peut donc l'écrire $b = h + \lambda a$ avec $h \in H$ et $\lambda \in \mathbb{R}$. Alors la suite de terme général $u_n = h + \lambda x_n$ est une suite d'éléments de H qui converge vers b.

En conclusion, H est dense dans E.

SF7.14 Utiliser certaines propriétés des fermés

Il s'agit d'utiliser les propriétés : une réunion d'une famille finie de fermés est un fermé, et une intersection d'une famille quelconque de fermés est un fermé.

Exemple

Soit E un espace préhilbertien, et A une partie non vide de E.

Montrer que $A^\perp$ est un fermé.

- *Considérons déjà le cas où A est réduit à un élément $\{a\}$.*
  Alors $\{a\}^\perp = \{x \in E \mid \langle a|x\rangle = 0\}$ est le noyau de la forme linéaire $\varphi : x \mapsto \langle a|x\rangle$.
  Or cette forme linéaire est continue puisque, d'après l'inégalité de Cauchy-Schwarz,
  $$\forall\, x \in E,\ |\varphi(x)| \leqslant k\,\|x\| \quad \text{avec } k = \|a\|\ ,$$
  et on reconnaît ici la caractérisation d'une application linéaire continue. Donc $\{a\}^\perp$, image réciproque de $\{0\}$ qui est un fermé de $\mathbb{R}$ par une application continue, est un fermé de E.
- *Dans le cas général, $A^\perp = \bigcap\limits_{a\in A} \{a\}^\perp$ est l'intersection de fermés, donc c'est un fermé.*

SF7.15 Utiliser les suites

Il s'agit ici de revenir simplement à la définition d'un compact à l'aide des suites extraites. Cela peut être aussi utilisé pour montrer qu'une partie n'est pas compacte.

Exemple

Soient A et B deux parties compactes d'un espace vectoriel normé E.
Montrer que $A+B$ est compacte.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de $A+B$. Il s'agit de montrer que l'on peut en extraire une suite convergente dans $A+B$.

Pour tout n, on a $x_n = a_n + b_n$, avec $a_n \in A$ et $b_n \in B$.
$(a_n)_{n\in\mathbb{N}}$ étant une suite de A compact, il existe $\varphi\colon \mathbb{N}\to\mathbb{N}$ strictement croissante telle que la suite $\left(a_{\varphi(n)}\right)_{n\in\mathbb{N}}$ converge vers $a \in A$.
La suite $\left(b_{\varphi(n)}\right)_{n\in\mathbb{N}}$ étant une suite de B compact, on peut aussi en extraire une suite convergente : il existe $\psi\colon \mathbb{N}\to\mathbb{N}$ strictement croissante telle que la suite $\left(b_{\varphi\circ\psi(n)}\right)_{n\in\mathbb{N}}$ converge vers $b \in B$.
Puisque la suite $\left(a_{\varphi\circ\psi(n)}\right)_{n\in\mathbb{N}}$ est extraite de la suite $\left(a_{\varphi(n)}\right)_{n\in\mathbb{N}}$, elle converge aussi vers a, et finalement la suite de terme général $x_{\varphi\circ\psi(n)}$ converge vers $a+b \in A+B$, ce qui établit le résultat voulu.

Exemple

Soit $E = \mathscr{C}([0\,;2\pi],\mathbb{C})$ muni de la norme de la convergence uniforme.
Montrer que la sphère unité $\mathcal{S}(0,1)$ de E n'est pas compacte.

Soit $(f_n)_{n\in\mathbb{N}}$ la suite d'éléments de E définis par : $f_n\colon t \mapsto \mathrm{e}^{\mathrm{i}nt}$. Pour tout $n\in\mathbb{N}$, $\|f_n\|_\infty = 1$ donc il s'agit d'une suite d'éléments de $\mathcal{S}(0,1)$.
Si n et m sont deux entiers distincts, on a, pour tout $t\in[0\,;2\pi]$:

$$\mathrm{e}^{\mathrm{i}nt} - \mathrm{e}^{\mathrm{i}mt} = \mathrm{e}^{\mathrm{i}\frac{n+m}{2}t}\left(2\mathrm{i}\sin\left(\frac{n-m}{2}t\right)\right),$$

donc $\|f_n - f_m\|_\infty = \max\limits_{t\in[0;2\pi]} \left|\mathrm{e}^{\mathrm{i}nt} - \mathrm{e}^{\mathrm{i}mt}\right| = 2$.

S'il existait une suite extraite $\left(f_{\varphi(n)}\right)_{n\in\mathbb{N}}$ convergente, on devrait avoir en particulier :

$$\lim_{n\to+\infty} \left\|f_{\varphi(n+1)} - f_{\varphi(n)}\right\|_\infty = 0,$$

ce qui est contredit le résultat précédent.
On a donc trouvé une suite de $\mathcal{S}(0,1)$ dont on ne peut extraire aucune suite convergente : $\mathcal{S}(0,1)$ n'est pas compacte.

✓ Cet exemple montre que, en dimension infinie, une partie fermée bornée n'est pas forcément compacte.

SF7.16 Utiliser une application continue bien choisie

Pour montrer qu'une partie est compacte, on peut montrer qu'il s'agit de l'image d'un compact par une application continue.

Exemple

Soient A et B deux parties compactes d'un espace vectoriel normé E.
Montrer que $A+B$ est compacte.

Considérons l'application $f : \begin{cases} E \times E & \longrightarrow & E \\ (x,y) & \longmapsto & x+y \end{cases}$, *où* $E \times E$ *est l'espace vectoriel normé produit, muni de la norme* $\|(x,y)\| = \max(\|x\|, \|y\|)$.

Cette application est continue, car somme des deux applications $(x,y) \mapsto x$ *et* $(x,y) \mapsto y$, *qui sont continues puisque lipschitziennes de rapport* 1.
L'ensemble $A \times B$ *est une partie compacte de* $E \times E$, *car produit de deux compacts (propriété du cours).*

La partie $A+B$, *image de* $A \times B$ *par* f, *est donc compacte.*

✓ *Cette démonstration peut sembler plus rapide que celle faite précédemment pour le même exemple. Mais en réalité, la démonstration du cours sur la compacité du produit* $A \times B$ *est en tous points similaire à celle que nous avons faite lors de cet exemple précédent.*

SF7.17 Utiliser le caractère fermé, ou fermé borné

Il s'agit d'utiliser le fait que toute partie fermée d'un compact est aussi un compact, ou que les parties compactes d'un espace vectoriel normé *de dimension finie* sont les parties fermées bornées.

Exemple

Montrer que l'ensemble $\mathcal{O}_n(\mathbb{R})$ des matrices orthogonales est une partie compacte de $\mathcal{M}_n(\mathbb{R})$.

Rappelons que : $\mathcal{O}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^tMM = \mathrm{I}_n\}$.
L'application $M \to {}^tMM$ *est continue sur* $\mathcal{M}_n(\mathbb{R})$, *comme composée d'applications continues : en effet, l'application* $M \mapsto {}^tM$ *est continue car linéaire sur un espace vectoriel de dimension finie, et l'application* $(M,N) \mapsto MN$ *est continue sur* $\mathcal{M}_n(\mathbb{R})^2$ *car multilinéaire sur des espaces vectoriels de dimensions finies.*

✓ *On peut aussi dire, plus simplement, que les applications coordonnées* $M \mapsto ({}^tMM)_{ij}$ *sont continues car polynomiales en les coefficients de* M.

L'ensemble $\mathcal{O}_n(\mathbb{R})$ *étant l'image réciproque du fermé* $\{\mathrm{I}_n\}$ *par cette application continue, c'est un fermé.*

D'autre part, considérons sur $\mathcal{M}_n(\mathbb{R})$ *la norme :* $\|M\|_\infty = \max_{1 \leqslant i,j \leqslant n} |m_{ij}|$ *(on peut bien sûr en utiliser une autre).*
Si M *est une matrice orthogonale, puisque ses vecteurs colonnes sont unitaires, ses coefficients sont inférieurs à* 1 *en valeur absolue. Donc pour toute* $M \in \mathcal{O}_n(\mathbb{R})$, $\|M\|_\infty \leqslant 1$, *ce qui montre que* $\mathcal{O}_n(\mathbb{R})$ *est une partie bornée.*

L'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ étant de dimension finie, on en déduit que $\mathcal{O}_n(\mathbb{R})$ est un compact.

■ Comment utiliser la compacité ?

SF7.18 Utiliser la propriété de Bolzano-Weierstrass

Il s'agit ici d'utiliser directement la caractérisation séquentielle d'un compact.

Exemple

Soient A une partie fermée et B une partie compacte d'un espace vectoriel normé E de dimension finie. On note $d(A,B) = \inf_{(x,y)\in A\times B} \|x-y\|$.

Montrer qu'il existe $a \in A$ et $B \in B$ tels que $d(A,B) = \|a-b\|$.

Notons $d = d(A,B)$. Par définition de la borne inférieure, pour tout $n \in \mathbb{N}$ il existe $(a_n,b_n) \in A \times B$ tel que :

$$d \leqslant \|a_n - b_n\| \leqslant d + \frac{1}{n+1} \quad (*)$$

B étant compacte, il existe une suite extraite $\left(b_{\varphi(n)}\right)_{n\in\mathbb{N}}$ qui converge vers $b \in B$. On a alors, pour tout entier n :

$$\left\|a_{\varphi(n)}\right\| \leqslant \left\|a_{\varphi(n)} - b_{\varphi(n)}\right\| + \left\|b_{\varphi(n)}\right\| \leqslant d + \frac{1}{n+1} + \left\|b_{\varphi(n)}\right\| ,$$

ce qui montre que la suite $\left(a_{\varphi(n)}\right)$ est bornée car la suite $\left(b_{\varphi(n)}\right)$ l'est. Puisque E est de dimension finie, cette suite admet une valeur d'adhérence, c'est-à-dire qu'on peut en extraire une suite $\left(a_{\varphi\circ\psi(n)}\right)$ qui converge vers $a \in E$. Mais A étant fermée, on a $a \in A$.
Les suites $\left(a_{\varphi\circ\psi(n)}\right)$ et $\left(b_{\varphi\circ\psi(n)}\right)$ convergent donc vers $a \in A$ et $b \in B$ respectivement, et en passant à la limite dans l'inégalité $(*)$ (appliquée à $a_{\varphi\circ\psi(n)}$ et $b_{\varphi\circ\psi(n)}$), on a $\|a-b\| = d$.

Exemple

Soit K une partie compacte d'un espace vectoriel normé, et $f : K \to K$ telle que :

$$\forall\,(x,y) \in K^2,\ \|f(x) - f(y)\| = \|x-y\| .$$

Montrer que f est une application continue bijective de K sur K.

- *f est continue car lipschitzienne de rapport 1.*
- *f est injective, car d'après l'hypothèse : $f(x) = f(y) \implies x = y$.*
- *Il reste à démontrer que $f(K) = K$.*
  Soit $x \in K$. On peut alors considérer la suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de K définie par récurrence par :

$$u_0 = x \quad \text{et} \quad \forall\, n \in \mathbb{N},\ u_{n+1} = f(u_n) .$$

  K étant compact, on peut en extraire une suite $\left(u_{\varphi(n)}\right)$ convergente dans K. En particulier, on a :

$$\lim_{n\to+\infty} \left\|u_{\varphi(n+1)} - u_{\varphi(n)}\right\| = 0 ,$$

donc si on se donne $\varepsilon > 0$ il existe $N \in \mathbb{N}$ tel que $\left\|u_{\varphi(N+1)} - u_{\varphi(N)}\right\| \leqslant \varepsilon$.
En posant $k = \varphi(N+1) - \varphi(N)$ ($k \in \mathbb{N}^*$), on a alors :

$$\begin{aligned}\left\|u_{\varphi(N)+k} - u_{\varphi(N)}\right\| &= \left\|f(u_{\varphi(N)+k-1}) - f(u_{\varphi(N)-1})\right\| \\ &= \left\|u_{\varphi(N)+k-1} - u_{\varphi(N)-1}\right\| ,\end{aligned}$$

d'où par récurrence $\left\|u_{\varphi(N)+k} - u_{\varphi(N)}\right\| = \|u_k - u_0\| = \|u_k - x\|$.
On a donc $\|u_k - x\| \leqslant \varepsilon$, et puisque $u_k = f(u_{k-1})$ appartient à $f(K)$, cela prouve que x appartient à $\overline{f(K)}$. On a donc $K \subset \overline{f(K)}$, mais, f étant continue, $f(K)$ est compact donc fermé et finalement, $f(K) = K$, c'est-à-dire f surjective.

SF7.19 **Dans un problème d'extremum**

Il s'agit certainement du théorème le plus important d'analyse : une fonction numérique continue sur un compact y est bornée et atteint ses bornes.

Exemple

Soit E un espace vectoriel normé, $F \neq E$ un sous-espace vectoriel de dimension finie de E. Montrer qu'il existe $x \in E$ tel que :

$$\|x\| = 1 \quad \text{et} \quad d(x,F) = 1 .$$

Puisque $F \neq E$, il existe $y \in E$ tel que $y \notin F$. Puisque F est de dimension finie, il est fermé, donc $d = d(x,F) > 0$ (car $d(x,F) = 0 \iff x \in \overline{F}$).
Soit B la boule fermée de centre x et de rayon $d+1$; il est clair que $d(x,F) = d(x,B \cap F)$. Or $B \cap F$ est un fermé borné dans F de dimension finie, donc c'est un compact, et puisque l'application $y \mapsto \|x-y\|$ est continue, sa borne inférieure sur ce compact est atteinte :

$$\exists\, y \in F \text{ tel que } \|x-y\| = d(x,B \cap F) = d(x,F) .$$

Posons enfin $z = \dfrac{x-y}{\|x-y\|}$.

- *on a évidemment $\|z\| = 1$;*
- *puisque $0_E \in F$, $d(z,F) \leqslant d(z,0_E) = \|z\| = 1$;*
- *pour tout $t \in F$,*

$$\begin{aligned}\|z-t\| = \left\|\frac{x-y}{\|x-y\|} - t\right\| &= \frac{1}{\|x-y\|}\left\|x - (y + \|x-y\|\, t)\right\| \\ &\geqslant \frac{1}{\|x-y\|} d(x,F) = 1 ,\end{aligned}$$

puisque $y + \|x-y\|\, t \in F$.

Finalement, $d(z,F) = 1$, donc z répond à la question.

Exemple

Soit K une partie compacte d'un espace vectoriel normé, et $f : K \to K$ telle que

$$\forall\, (x,y) \in K^2,\ x \neq y \implies \|f(x) - f(y)\| < \|x-y\| .$$

Montrer que f possède un unique point fixe $a \in K$.

- *L'unicité est facile, car si a et b étaient deux points fixes distincts de f, on aurait $\|f(a)-f(b)\| < \|a-b\|$, soit $\|a-b\| < \|a-b\|$, c'est impossible.*
- *Raisonnons par l'absurde. Si f n'avait pas de point fixe, on aurait $f(x) \neq x$ pour tout $x \in K$.*
  Soit $\varphi : x \in K \mapsto d(x,f(x))$. Puisque la distance est une application continue sur E^2 et que f est continue sur K, φ est une application continue de K dans $\mathbb{R}_+$.
  K étant compact, φ est bornée et atteint ses bornes sur K, et en particulier :
  $$\exists\, y \in K \text{ tel que } \varphi(y) = \min_{x\in K} \varphi(x)\,.$$
  Puisque $y \neq f(y)$ on a alors :
  $$\varphi(f(y)) = \|f(f(y)) - f(y)\| < \|f(y) - y\| = \varphi(y)\,,$$
  d'où une contradiction avec la définition de y.
  On a donc prouvé par l'absurde que f admet un point fixe dans K.

■ Comment montrer et utiliser qu'une partie est connexe par arcs ?

SF7.20 Construire un chemin

Exemple

Soient A et B deux parties connexes par arcs d'un espace vectoriel normé E.
Montrer que $A+B$ est connexe par arcs.

Soient x, $y \in A+B$: $x = a+b$, $y = a'+b'$, avec $(a,a') \in A^2$ et $(b,b') \in B^2$.
A étant connexe par arcs, il existe une application continue $\gamma_A : [0\,;1] \to A$ telle que :
$$\gamma_A(0) = a,\ \gamma_A(1) = a' \text{ et } \forall\, t \in [0\,;1],\ \gamma_A(t) \in A\,.$$
De même, B étant connexe par arcs, il existe une application continue $\gamma_B : [0\,;1] \to B$ telle que :
$$\gamma_B(0) = b,\ \gamma_B(1) = b' \text{ et } \forall\, t \in [0\,;1],\ \gamma_B(t) \in B\,.$$
Posons alors, pour tout $t \in [0\,;1]$, $\varphi(t) = \gamma_A(t)+\gamma_B(t)$. Alors φ est continue comme somme d'applications continues, et il est facile de vérifier qu'il s'agit d'un chemin reliant $a+a'$ à $b+b'$, et entièrement contenu dans $A+B$.
On peut donc relier tous les points x et y de $A+B$ par un chemin continu contenu dans $A+B$: $A+B$ est connexe par arcs.

Exemple

1) Montrer que si une partie non vide A d'un espace vectoriel normé est étoilée par rapport à l'un de ses points, elle est connexe par arcs.

2) Montrer que l'ensemble des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ est connexe par arcs.

3) Montrer que l'ensemble des matrices nilpotentes de $\mathcal{M}_n(\mathbb{R})$ est connexe par arcs.

1) *Supposons A étoilé par rapport à a. Soient x et y deux points de A. On sait que le segment $[a\,;x]$ et le segment $[a\,;y]$ sont inclus dans A.*
On peut alors considérer le chemin $\gamma\colon [0\,;1]\to E$ défini par :

$$\gamma(t)=\begin{cases} a+(1-2t)(x-a) & \text{si } 0\leqslant t\leqslant \dfrac{1}{2}\\ a+(2t-1)(y-a) & \text{si } \dfrac{1}{2}<t\leqslant 1\end{cases}$$

(on parcourt d'abord le segment de x à a puis le segment de a à y).
Il est alors facile de vérifier que γ est continue (faire la limite à droite et à gauche en $\frac{1}{2}$), que $\gamma(0)=x$, $\gamma(1)=y$ et que pour tout t, $\gamma(t)\in A$ (car pour $t\in\left[0\,;\frac{1}{2}\right]$, $\gamma(t)$ appartient au segment $[x\,;a]$, et au segment $[a\,;y]$ pour $t\in\left[\frac{1}{2}\,;1\right]$).
On en déduit que A est connexe par arcs.

2) *Montrons que l'ensemble $\mathcal{D}_n$ des matrices diagonalisables de $\mathcal{M}_n(\mathbb{R})$ est étoilé par rapport à I_n, ce qui impliquera qu'il est connexe par arcs.*
En effet, soit $M\in\mathcal{D}_n$. Il existe $P\in\mathrm{GL}_n(\mathbb{R})$ et D diagonale telles que $M=PDP^{-1}$. Pour tout $t\in[0\,;1]$, on a :

$$tM+(1-t)\mathrm{I}_n=P\big(tD+(1-t)\mathrm{I}_n\big)P^{-1}\,,$$

et puisque la matrice $tD+(1-t)\mathrm{I}_n$ est diagonale, les matrices $tM+(1-t)\mathrm{I}_n$ sont diagonalisables, c'est-à-dire que le segment $[\mathrm{I}_n\,;M]$ est inclus dans $\mathcal{D}_n$, ce qui achève la démonstration.

3) *Puisque si N est nilpotente, il en est de même de toute matrice tN avec $t\in[0\,;1]$, l'ensemble $\mathcal{N}$ des matrices nilpotentes est étoilé par rapport à 0_n, donc est connexe par arcs.*

SF7.21 Utiliser une application continue

Exemple

Montrer que $\mathrm{GL}_n(\mathbb{R})$ n'est pas une partie connexe par arcs de $\mathcal{M}_n(\mathbb{R})$.

On sait que l'application $M\mapsto\det M$ est continue sur $\mathcal{M}_n(\mathbb{R})$ (car, par exemple, c'est une fonction polynomiale des coefficients).
Si $\mathrm{GL}_n(\mathbb{R})$ était connexe par arcs, son image par cette application serait une partie connexe de $\mathbb{R}$ c'est-à-dire un intervalle. Or cette image est égale à $\mathbb{R}^*$, d'où la contradiction.

Exemple

Montrer que le groupe spécial orthogonal $\mathcal{O}_2^+(\mathbb{R})$ est connexe par arcs.

On sait que : $\mathcal{O}_2^+(\mathbb{R})=\left\{\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\ \theta\in\mathbb{R}\right\}$.

C'est donc l'image de $\mathbb{R}$ par l'application $\begin{cases}\mathbb{R} & \longrightarrow & \mathcal{M}_2(\mathbb{R})\\ \theta & \longmapsto & \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\end{cases}$,

qui est continue car ses fonctions coordonnées le sont.
$\mathbb{R}$ étant connexe par arcs, il en est de même de son image.

SF7.22 Utiliser la connexité par arcs

On utilise le plus souvent la connexité par arcs avec le théorème suivant (déjà utilisé ci-dessus) :

> L'image d'une partie connexe par arcs par une application continue est une partie connexe par arcs.

Cela permet d'obtenir de façon très élégante certains résultats d'analyse réelle.

Exemple

Montrer qu'il n'existe pas d'application continue de $\mathbb{R}$ dans $\mathbb{R}$ telle que l'image de tout rationnel soit un irrationnel et que l'image de tout rationnel soit un irrationnel.

Supposons qu'une telle fonction f existe, et considérons alors l'application $g\colon x \in \mathbb{R} \mapsto f(x) + x$.
Alors g est continue sur $\mathbb{R}$, et $g(\mathbb{R}) \subset \mathbb{R}\setminus\mathbb{Q}$. L'image de la partie $\mathbb{R}$ connexe par arcs n'est donc pas connexe par arcs (rappelons que les parties connexes par arcs de $\mathbb{R}$ sont les intervalles), c'est impossible.

Exemple

Soit I in intervalle de $\mathbb{R}$, et $f\colon I \to \mathbb{R}$ une application continue et injective.
Montrer que f est strictement monotone.

Considérons la partie de $\mathbb{R}^2$ suivante : $\Delta = \{(x,y) \in I^2 \mid x < y\}$.
Δ est une partie connexe par arcs de $\mathbb{R}^2$ (car c'est aussi une partie convexe).
Soit alors $F : \begin{cases} \Delta & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f(x) - f(y) \end{cases}$. F est continue car f l'est, donc $f(\Delta)$ est une partie connexe par arcs de $\mathbb{R}$. Or $f(\Delta) \subset \mathbb{R}^*$ puisque f est injective. Il en résulte que l'on a :

- *soit $f(\Delta) \subset \mathbb{R}_+^*$, c'est-à-dire $f(x) > f(y)$ dès que $x < y$, et f est strictement décroissante ;*
- *soit $f(\Delta) \subset \mathbb{R}_-^*$, et f est strictement croissante.*

■ Comment montrer qu'une application linéaire est continue ?

SF7.23 Montrer qu'elle est lipschitzienne

Pour montrer qu'une application linéaire entre deux espaces vectoriels normés est continue, on ne revient *jamais* à la définition de la continuité en calculant des limites ! On utilise exclusivement la caractérisation rappelée ci-dessous (cf. **[S7.8]**).

> Une application $u \in \mathscr{L}(E,F)$ est continue si et seulement si il existe $k \in \mathbb{R}_+$ tel que, pour tout x de E, on ait $\|u(x)\|_F \leqslant k\,\|x\|_E$.

Nous avons d'ailleurs déjà utilisé cette propriété dans un exemple de **[SF7.14]** page 379.

Exemple

Pour tout polynôme $P = \sum_{k=0}^{n} a_k X^k$ de $\mathbb{R}[X]$, on pose $N(P) = \sum_{k=0}^{n} |a_k|$.

1) Montrer que N est une norme sur $\mathbb{R}[X]$.

2) On considère les applications u et v définies sur $\mathbb{R}[X]$ par

$$\forall\, P \in \mathbb{R}[X],\ u(P) = P(0) \quad \text{et} \quad v(P) = P'\,.$$

Étudier la continuité de u et de v sur l'espace vectoriel normé (E,N).

1) *Le fait que N est une norme ne doit pas poser problème.*

2) • *Déjà, le fait que u et v sont des applications linéaires est immédiat.*

• *Pour tout $P = \sum_{k=0}^{n} a_k X^k$ on a :*

$$|u(P)| = |P(0)| = |a_0| \leqslant N(P)\,,$$

donc u est continue, d'après la caractérisation de la continuité pour une application linéaire.

• *Considérons la suite de polynômes $(P_n)_{n\in\mathbb{N}}$ donnés par $P_n = X^n$.*
Alors $N(P_n) = 1$ et $N(v(P_n)) = N(nX^{n-1}) = n$ (si $n \geqslant 1$).
La suite $(P_n)_{n\in\mathbb{N}}$ est donc bornée pour N alors que la suite $\big(v(P_n)\big)_{n\in\mathbb{N}}$ ne l'est pas, donc v n'est pas continue pour la norme N.

SF7.24 Utiliser la dimension

On sait que si l'espace vectoriel normé E est de dimension finie, et F un espace vectoriel normé quelconque, toute application linéaire $u \in \mathscr{L}(E,F)$ est continue.

Il n'y a donc rien à faire dans ce cas, mais cette propriété est très utile. Nous nous en sommes déjà servi, et voici un autre exemple important d'application.

Exemple

Soit $A \in \mathcal{M}_p(\mathbb{K})$. On suppose que la suite $(A^n)_{n\in\mathbb{N}}$ converge vers une certaine matrice P.
Montrer que P et A commutent, et que P est une matrice de projection.

• *On a : $A^{n+1} = A^n.A = A.A^n$. Il s'agit de* justifier rigoureusement *le passage à la limite dans ces égalités.*

– *La suite (A^{n+1}) est une suite extraite de la suite (A^n), donc $\lim\limits_{n\to+\infty} A^{n+1} = \lim\limits_{n\to+\infty} A^n = P$.*

– *L'application $u\colon M \mapsto MA$ est un endomorphisme de $\mathcal{M}_p(\mathbb{K})$ (vérification facile). $\mathcal{M}_p(\mathbb{K})$ étant de dimension finie, u est continue et, par la caractérisation séquentielle de la continuité, on a :*

$$\lim_{n\to+\infty} u(A^n) = u\left(\lim_{n\to+\infty} A^n\right),$$

c'est-à-dire : $\lim\limits_{n\to+\infty} A^{n+1} = u(P) = PA\,.$

– *En considérant l'endomorphisme* $M \mapsto AM$ *on démontre de même que :* $\lim\limits_{n\to+\infty} A^{n+1} = AP$.

La juxtaposition de ces trois résultats donne : $P = PA = AP$.

- *La suite* (A^{2n}) *est une suite extraite de la suite* (A^n) *donc* $\lim\limits_{n\to+\infty} A^{2n} = P$.

D'autre part, l'application $B : \begin{cases} \mathcal{M}_p(\mathbb{K})^2 & \longrightarrow & \mathcal{M}_p(\mathbb{K}) \\ (M,N) & \longmapsto & MN \end{cases}$ *est bilinéaire.*

$\mathcal{M}_p(\mathbb{K})$ *étant un espace vectoriel de dimension finie,* B *est continue et, par la caractérisation séquentielle de la continuité, on a :*

$$\lim_{n\to+\infty} B(A^n,A^n) = B\Big(\lim_{n\to+\infty} A^n, \lim_{n\to+\infty} A^n\Big) = B(P,P) = P^2 .$$

Puisque $B(A^n,A^n) = A^{2n}$, *on en déduit* $P^2 = P$, *c'est-à-dire que* P *est la matrice d'un projecteur.*

■ Comment étudier la continuité d'une application ?

Le cas des applications linéaires ayant été traité plus haut, nous considérons ici des applications définies sur une partie D d'un espace vectoriel normé E et à valeurs dans un espace vectoriel normé F.

> ✓ Attention ! Le critère $\|u(x)\|_F \leqslant k\,\|x\|_E$ pour montrer la continuité n'est valable que si u est une *application linéaire*. Ne l'employez donc pas pour une application quelconque !

SF7.25 Utiliser le caractère lipschitzien

Parfois, on peut démontrer simplement la continuité d'une application en vérifiant qu'elle est lipschitzienne, voire localement lipschitzienne (rappelons que la réciproque de cette propriété est fausse).

Exemple

Soit A une partie non vide d'un espace vectoriel normé E. Montrer que l'application $x \mapsto d(x,A)$ est continue.

Pour tous x, $y \in E$ *et tout* $a \in A$, *l'inégalité triangulaire donne :*

$$\|x-a\| \leqslant \|x-y\| + \|y-a\| .$$

$d(x,A) = \inf\limits_{a\in A} \|x-a\|$ *donc par définition de la borne inférieure :*

$\forall\, a \in A,\ d(x,A) \leqslant \|x-y\| + \|y-a\|$ *soit* $d(x,A) - \|x-y\| \leqslant \|y-a\|$,

d'où l'on tire, encore par définition de la borne inférieure :

$$d(x,A) - \|x-y\| \leqslant d(y,A) .$$

Ainsi $d(x,A) - d(y,A) \leqslant \|x-y\|$ *puis, en échangeant les rôles de* x *et* y *:*

$$|d(x,A) - d(y,A)| \leqslant \|x-y\| .$$

L'application $x \mapsto d(x,A)$ *est donc lipschitzienne de rapport* 1 *et, par suite, elle est continue.*

✓ *Remarquez dans cet exemple comment nous avons utilisé de façon rigoureuse la définition de la borne inférieure, en procédant en deux temps. Il eût été incorrect d'écrire, de façon approximative, « on passe à la borne inférieure dans l'inégalité* $\|x-a\| \leqslant \|x-y\| + \|y-a\|$ *».*

SF7.26 Utiliser les théorèmes généraux

Il s'agit de la méthode à utiliser systématiquement lorsqu'il n'y a pas de problème : on essaye d'utiliser les théorèmes relatifs aux opérations sur les fonctions continues.

Exemple

Soit E un espace vectoriel normé.
Montrer que l'application $f : x \mapsto \dfrac{x}{1+\|x\|^2}$ est continue de E dans E.

L'application $\mathrm{Id}_E : x \mapsto x$ *est évidemment continue.*
L'application $x \mapsto \|x\|$ *est continue de* E *dans* $\mathbb{R}$ *(elle est lipschitzienne de rapport* 1*) donc d'après les théorèmes sur les fonctions à valeurs réelles il en est de même de* $\varphi : x \mapsto \dfrac{1}{1+\|x\|^2}$.
Donc $f = \varphi.\mathrm{Id}_E$ *est continue, d'après les théorèmes du cours.*

SF7.27 Revenir à la définition

Il ne s'agit pas ici de manipuler des ε, mais d'utiliser les limites : une fonction f est continue en a si $\lim\limits_{x\to a} f(x) = f(a)$. On reviendra donc aux calculs de limites aux points qui posent problème.

Exemple

Étudier la continuité de l'application $f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \lfloor x \rfloor + \left(x - \lfloor x \rfloor\right)^2 \end{cases}$.

- *Soit* $n \in \mathbb{Z}$. *Pour tout* $x \in]n\,;n+1[$ *on a* $f(x) = n+(x-n)^2$ *donc d'après les théorèmes généraux,* f *est continue sur* $]n\,;n+1[$.
- f *est donc continue sur* $\mathbb{R}\setminus\mathbb{Z}$, *et il reste à étudier sa continuité aux points entiers.*
  Soit donc $n \in \mathbb{Z}$. *D'après l'expression de* f *ci-dessus, on a :*
  $$\lim_{\substack{x\to n\\ x>n}} f(x) = \lim_{x\to n} \left(n + (x-n)^2\right) = n\,.$$
  Lorsque x *tend vers* n *par valeurs inférieures, on peut supposer* $x \in]n-1\,;n[$ *donc dans ce cas* $f(x) = (n-1) + (x-(n-1))^2$ *et*
  $$\lim_{\substack{x\to n\\ x<n}} f(x) = \lim_{x\to n} \left((n-1) + (x-(n-1))^2\right) = (n-1)+1 = n\,.$$
  Ainsi $\lim\limits_{\substack{x\to n\\ x>n}} f(x) = \lim\limits_{\substack{x\to n\\ x<n}} f(x) = n = f(n)$, *donc* f *est continue en* n.
- *En conclusion,* f *est continue sur* $\mathbb{R}$.

SF7.28 Utiliser la caractérisation séquentielle

Le théorème relatif à la caractérisation séquentielle de la continuité est ici utilisé pour montrer par contraposition qu'une fonction n'est pas continue en un point (cf. aussi l'exemple de **[SF7.23]** page 387).

Exemple

Monter que l'application $f\colon x \in \mathbb{R} \longmapsto \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{si } x \notin \mathbb{Q} \end{cases}$ est discontinue en tout point.

- *Soit $a \in \mathbb{R} \setminus \mathbb{Q}$. On sait qu'il existe une suite de rationnels $(x_n)_{n\in\mathbb{N}}$ qui converge vers a.*
  On a alors : $\forall n \in \mathbb{N},\ f(x_n) = 1$, donc $\lim\limits_{n\to+\infty} f(x_n) = 1 \neq f(a) = 0$.
  Donc f n'est pas continue en a.
- *On montre de la même façon que f est discontinue en tout point $a \in \mathbb{Q}$, en utilisant le fait que dans tout voisinage de a il existe des irrationnels, donc que a est limite d'une suite d'irrationnels.*

■ Comment étudier l'uniforme continuité d'une application ?

SF7.29 Utiliser le caractère lipschitzien

On sait que toute application lipschitzienne est uniformément continue (la réciproque de cette propriété est fausse). Dans beaucoup de cas, cette propriété permet de démontrer facilement l'uniforme continuité (et donc la continuité). Nous avons déjà donné un exemple de cette méthode dans un exemple récent page 388, en voici un autre.

Exemple

Soit $E = \mathscr{C}\big([0\,;1],\mathbb{R}\big)$, muni de la norme $\|\cdot\|_\infty$.

Montrer que l'application $\varphi\colon \begin{cases} E & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \inf\limits_{t\in[0;1]} f(t) \end{cases}$ est uniformément continue.

Soient f, $g \in E$. Pour tout $t \in [0\,;1]$ on a :

$$f(t) = g(t) + f(t) - g(t) \leqslant g(t) + |f(t) - g(t)| \leqslant g(t) + \|f - g\|_\infty .$$

On en déduit $\varphi(f) \leqslant g(t) + \|f - g\|_\infty$, soit $\varphi(f) - \|f - g\|_\infty \leqslant g(t)$.

Cela étant vrai pour tout $t \in [0\,;1]$, on en déduit $\varphi(f) - \|f - g\|_\infty \leqslant \varphi(g)$, soit $\varphi(f) - \varphi(g) \leqslant \|f - g\|_\infty$.

En échangeant les rôles de f et g on a de même $\varphi(g) - \varphi(f) \leqslant \|g - f\|_\infty$ donc finalement :

$$|\varphi(f) - \varphi(g)| \leqslant \|f - g\|_\infty .$$

L'application φ est donc lipschitzienne de rapport 1 et par suite, elle est uniformément continue.

SF7.30 Revenir directement à la définition

Exemple

Montrer que la fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}_+$.

Un calcul simple permet de montrer que :

$$\forall\,(x,y) \in (\mathbb{R}_+)^2,\ \left|\sqrt{x}-\sqrt{y}\right| \leqslant \sqrt{|x-y|}$$

(il suffit d'élever au carré, puis de distinguer les cas $x \leqslant y$ ou $y \leqslant x$).
Soit alors $\varepsilon > 0$, et posons $\alpha = \varepsilon^2$. L'inégalité précédente montre directement que :

$$\forall\,(x,y) \in (\mathbb{R}_+)^2,\ |x-y| \leqslant \alpha \implies \left|\sqrt{x}-\sqrt{y}\right| \leqslant \varepsilon\,,$$

donc la fonction $\sqrt{\ }$ est uniformément continue sur $\mathbb{R}_+$ par définition.

SF7.31 Raisonner par l'absurde

Montrer par l'absurde qu'une application est uniformément continue est la méthode utilisée dans le cours pour démontrer le théorème de Heine. Voici un autre exemple d'application.

Exemple

Soit E le $\mathbb{R}$-espace vectoriel des fonctions continues bornées de $\mathbb{R}$ dans $\mathbb{R}$, muni de la norme $\|\cdot\|_\infty$.

Pour tout $f \in E$ et tout $a \in \mathbb{R}$, on note f_a l'application $f_a : x \mapsto f(x+a)$.

1) Montrer que si $\{f_a,\ a \in R\}$ est une partie compacte de E, alors f est uniformément continue sur $\mathbb{R}$.

2) En considérant $f : x \in \mathbb{R} \mapsto \dfrac{1}{1+x^2}$, montrer que la réciproque de cette propriété est fausse.

1) *Notons $A = \{f_a, a \in \mathbb{R}\}$.*
Supposons par l'absurde que f n'est pas uniformément continue, c'est-à-dire qu'il existe $\varepsilon > 0$ (fixé pour la suite) tel que

$$\forall\,\alpha > 0,\ \exists\, x,y \in \mathbb{R} \text{ tels que } |y-x| \leqslant \alpha \text{ et } |f(y)-f(x)| > \varepsilon\,.$$

En appliquant cette propriété avec $\alpha = \dfrac{1}{n+1}$, on construit deux suites réelles $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ telles que $|y_n - x_n| \leqslant \dfrac{1}{n+1}$ et $|f(y_n)-f(x_n)| > \varepsilon$.
Posons alors $a_n = y_n - x_n$. La suite $\left(f_{a_n}\right)$ étant une suite d'éléments de A compact, on peut en extraire une suite $\left(f_{a_{\varphi(n)}}\right)$ qui converge vers $f_a \in A$ au sens de la norme $\|\cdot\|_\infty$, c'est-à-dire que la suite de fonctions $\left(f_{a_{\varphi(n)}}\right)$ converge uniformément vers f_a sur $\mathbb{R}$.
Elle converge donc aussi simplement, c'est-à-dire que pour tout $x \in \mathbb{R}$, $\lim\limits_{n\to+\infty} f(x+a_{\varphi(n)}) = f(x+a)$. Mais la fonction f est continue, et la suite $\left(a_{\varphi(n)}\right)$ tend vers 0, donc $\lim\limits_{n\to+\infty} f(x+a_{\varphi(n)}) = f(x)$, soit $f(x+a) = f(x)$ pour tout x, c'est-à-dire $f_a = f$.

Ainsi la suite $\left(f_{a_{\varphi(n)}}\right)$ *converge vers* f : $\lim\limits_{n\to+\infty} \left\|f_{a_{\varphi(n)}} - f\right\|_\infty = 0$.
Or pour tout n :

$$\varepsilon < |f(y_n) - f(x_n)| = |f_{a_n}(x_n) - f(x_n)| \leqslant \|f_{a_n} - f\|_\infty ,$$

d'où la contradiction.

2) *L'application* $f\colon x \in \mathbb{R} \mapsto \dfrac{1}{1+x^2}$ *est lispschitzienne de rapport* 2 *(car* $\|f'\|_\infty^{\mathbb{R}} = 2$*), donc elle est uniformément continue sur* $\mathbb{R}$.
Si A *était compacte, on pourrait extraire de la suite des fonctions* $f_n\colon x \mapsto \dfrac{1}{1+(x+n)^2}$ *une suite* $\left(f_{\varphi(n)}\right)$ *qui serait uniformément convergente sur* $\mathbb{R}$ *vers une fonction* f_a *avec* $a \in \mathbb{R}$.
Comme la convergence uniforme entraîne la convergence simple, on devrait donc avoir, pour tout $x \in \mathbb{R}$:

$$\lim_{n\to+\infty} \frac{1}{1+(x+\varphi(n))^2} = \frac{1}{1+(x+a)^2} ,$$

ce qui est impossible puisque $\lim\limits_{n\to+\infty} \varphi(n) = +\infty$.

La méthode du raisonnement par l'absurde est aussi utile pour démontrer qu'une application n'est *pas* uniformément continue.

Exemple

Montrer que la fonction $f\colon x \mapsto \sin(x^2)$ n'est pas uniformément continue sur $\mathbb{R}$.

Si f *était uniformément continue, on aurait, par définition :*

$$\forall\, \varepsilon > 0,\ \exists\, \alpha > 0,\ \forall\, (x,y) \in \mathbb{R}^2,\ |x-y| \leqslant \alpha \implies |f(x) - f(y)| \leqslant \varepsilon .$$

Si pour tout $n \in \mathbb{N}$ *on pose* $x_n = \sqrt{2n\pi}$ *et* $y_n = \sqrt{2n\pi + \frac{\pi}{2}}$, *on a :*

$$y_n - x_n = \sqrt{2n\pi + \tfrac{\pi}{2}} - \sqrt{2n\pi} = \frac{\frac{\pi}{2}}{\sqrt{2n\pi + \frac{\pi}{2}} + \sqrt{2n\pi}}$$

donc $\lim\limits_{n\to+\infty} |x_n - y_n| = 0$ *mais* $|f(x_n) - f(y_n)| = 1$ *pour tout* n, *ce qui contredit la définition précédente.*
Cela montre que f *n'est pas uniformément continue sur* $\mathbb{R}$.

✓ *On a ainsi trouvé un exemple de fonction continue bornée mais non uniformément continue.*

SF7.32 Utiliser le théorème de Heine

Le théorème de Heine est utilisé dans la démonstration de nombreux théorèmes importants, concernant par exemple l'approximation uniforme d'une fonction continue sur un segment. Il peut aussi être utilisé pour démontrer facilement l'uniforme continuité d'une fonction, même si elle n'est pas définie sur une partie compacte.

Exemple

Soit f une fonction continue de $\mathbb{R}_+$ dans $\mathbb{R}$ telle que $\lim\limits_{x\to+\infty} f(x) = \ell$ existe.
Montrer que f est uniformément continue sur $\mathbb{R}_+$.

Soit $\varepsilon > 0$. Par définition de la limite :

$$\exists\, A \in \mathbb{R}_+,\ x \geqslant A \implies |f(x) - \ell| \leqslant \frac{\varepsilon}{4}\cdot$$

On a donc, grâce à l'inégalité triangulaire :

$$\forall\,(x,y) \in [A\,;+\infty[^2,\ |f(x) - f(y)| \leqslant \frac{\varepsilon}{2}\cdot$$

f étant continue sur l'intervalle compact $[0\,;A]$, elle y est uniformément continue, donc il existe $\alpha > 0$ tel que :

$$\forall\,(x,y) \in [0\,;A]^2,\ |x-y| \leqslant \alpha \implies |f(x) - f(y)| \leqslant \frac{\varepsilon}{2}\cdot$$

Si maintenant x et y sont tels que $0 \leqslant x \leqslant A \leqslant y$ et $|x-y| \leqslant \alpha$, on a forcément $|x - A| \leqslant \alpha$ donc $|f(x) - f(A)| \leqslant \frac{\varepsilon}{2}$, et aussi, puisque $y \geqslant A$, $|f(A) - f(y)| \leqslant \frac{\varepsilon}{2}$, donc finalement $|f(x) - f(y)| \leqslant \varepsilon$ par l'inégalité triangulaire.

Finalement on a bien, pour tous x, $y \in \mathbb{R}_+$:

$$|x-y| \leqslant \alpha \implies |f(x) - f(y)| \leqslant \varepsilon\,,$$

ce qui prouve que f est uniformément continue sur $\mathbb{R}_+$.

■ Comment utiliser la structure d'espace vectoriel normé de $\mathcal{M}_n(\mathbb{K})$?

De fort jolies propriétés peuvent être obtenues sur les espaces de matrices en utilisant conjointement des méthodes algébriques (principalement la réduction) et des méthodes d'analyse. Nous avons déjà eu l'occasion d'étudier des suites de matrices ou des propriétés topologiques de certains ensembles de matrices dans de nombreux exemples précédents, tellement le sujet est riche. Nous présentons ci-après d'autres exemples.

SF7.33 Utiliser les normes

Bien que les toutes les normes sur $\mathcal{M}_n(\mathbb{K})$ soient équivalentes, certaines possèdent des propriétés plus intéressantes que d'autres, en particulier les normes d'algèbre, utilisées par exemple pour l'étude de séries matricielles. Rappelons qu'une norme d'algèbre sur $\mathcal{M}_n(\mathbb{K})$ est une norme qui vérifie en outre :

$$\forall\, A,\, B \in \mathcal{M}_n(\mathbb{K}),\ \|AB\| \leqslant \|A\|\,\|B\|\ .$$

Les normes définies ci-dessous pour $A = (a_{ij})_{1\leqslant i,j\leqslant n}$ sont des normes d'algèbre :

- $\|A\| = \displaystyle\sum_{1\leqslant i,j\leqslant n} |a_{ij}|$;
- $\|A\| = \left(\displaystyle\sum_{1\leqslant i,j\leqslant n} |a_{ij}|^2\right)^{1/2}$ (voir **[SF5.3]** page 336) ;
- $\|A\| = \displaystyle\max_{1\leqslant i\leqslant n}\left(\sum_{j=1}^{n} |a_{ij}|\right)$ (voir exercice **7.2**) et $\|A\| = \displaystyle\max_{1\leqslant j\leqslant n}\left(\sum_{i=1}^{n} |a_{ij}|\right)$.

Exemple

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, on note $\rho(A) = \max\{|\lambda|\,,\ \lambda \in \mathrm{Sp}_{\mathbb{C}}(A)\}$ le rayon spectral de A. Prouver que si N est une norme d'algèbre sur $\mathcal{M}_n(\mathbb{C})$, on a :

$$\forall\, A \in \mathcal{M}_n(\mathbb{C}),\ \rho(A) \leqslant N(A)\,.$$

Soit $\lambda \in \mathrm{Sp}(A)$ tel que $|\lambda| = \rho(A)$.
Il existe $X \in \mathcal{M}_{n,1}(\mathbb{C})$, $X \neq 0$, tel que $AX = \lambda X$.
Considérons alors la matrice $M = \begin{bmatrix} X & X & \dots & X \end{bmatrix}$ dont toutes les colonnes sont égales à X. Alors $AM = \begin{bmatrix} AX & AX & \dots & AX \end{bmatrix} = \lambda M$, et la relation $N(AM) \leqslant N(A)N(M)$ donne $|\lambda| \leqslant N(A)$ puisque $N(M) \neq 0$.

SF7.34 Utiliser la densité

Exemple

1) Montrer que $\mathrm{GL}_n(\mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$.

2) Applications : en commençant par étudier le cas de matrices inversibles, démontrer les propriétés suivantes :

a) Si A, $B \in \mathcal{M}_n(\mathbb{K})$, le polynôme caractéristique de AB est égal à celui de BA.

b) Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $N \in \mathcal{M}_n(\mathbb{K})$ nilpotente telles que $AN = NA$. Montrer que $\det(A + N) = \det A$.

1) • *En* **[SF7.10]** *page 377, nous avons déjà démontré que* $\mathrm{GL}_n(\mathbb{K})$ *est un ouvert de* $\mathcal{M}_n(\mathbb{K})$.

• *Pour montrer qu'il est dense dans* $\mathcal{M}_n(\mathbb{K})$, *il suffit de montrer que toute matrice* $A \in \mathcal{M}_n(\mathbb{K})$ *est limite d'une suite de matrices inversibles.*

Notons $\rho = +\infty$ *si* A *n'a pas de valeur propre non nulle, et* $\rho = \min\{|\lambda|\,,\ \lambda \in \mathrm{Sp}(A),\ \lambda \neq 0\}$ *sinon ; soit* n_0 *un entier tel que* $0 < \dfrac{1}{n_0} < \rho$. *Pour tout entier* $p \geqslant n_0$, $\dfrac{1}{p}$ *n'est pas valeur propre de* A, *donc la matrice* $A_p = A - \dfrac{1}{p}\mathrm{I}_n$ *est inversible. Il est clair que* $\lim\limits_{p\to+\infty} A_p = A$; *on a donc trouvé une suite de matrices inversibles qui converge vers* A.

2) a) • *Supposons d'abord* A *inversible. Alors* $BA = A^{-1}(AB)A$ *donc* BA *est semblable à* AB *et par suite a même polynôme caractéristique.*

• *L'application* $\chi : \left\{ \begin{array}{ccc} \mathcal{M}_n(\mathbb{K}) & \longrightarrow & \mathbb{K}[X] \\ M & \longmapsto & \chi_M \end{array} \right.$ *qui à toute matrice* M *associe son polynôme caractéristique est continue. Pour le démontrer, il suffit de remarquer que chaque application coordonnée (c'est-à-dire celle qui à* M *associe le coefficient de* X^k*) est continue puisque c'est une somme de déterminants extraits de* M, *donc est une fonction polynomiale des coefficients de* M.

L'application qui à deux matrices A *et* B *associe leur produit est également continue (car bilinéaire en dimension finie), donc l'application* $A \mapsto \chi_{AB} - \chi_{BA}$ *est continue de* $\mathcal{M}_n(\mathbb{K})$ *dans* $\mathbb{K}[X]$. *Puisqu'elle est nulle*

sur $\mathrm{GL}_n(\mathbb{K})$ *qui est une partie dense, elle est nulle sur* $\mathcal{M}_n(\mathbb{K})$*, ce qui établit le résultat.*

b) • *Supposons d'abord* $A = \mathrm{I}_n$*.* N *étant nilpotente, elle est semblable à une matrice triangulaire supérieure* T *à éléments diagonaux nuls (cf.* **[S4.13]***). On a alors* $\det(\mathrm{I}_N + N) = \det(\mathrm{I}_n + T) = 1$*, ce qui établit le résultat dans ce cas.*

• *Supposons maintenant* A *inversible. Alors* $A + N = A(\mathrm{I}_n + A^{-1}N)$*. Puisque* A *et* N *commutent, il en est de même de* A^{-1} *et de* N*, donc* $(A^{-1}N)^n = A^{-n}N^n = 0_n$ *et* $A^{-1}N$ *est nilpotente.*

D'après le point précédent, $\det(A+N) = \det A \cdot \det(\mathrm{I}_n + A^{-1}N) = \det A$*, et le résultat est démontré dans le cas* A *inversible.*

• *Dans le cas général, on écrit* $A = \lim\limits_{p\to+\infty} \left(A - \dfrac{1}{p}\mathrm{I}_n\right)$ *où les matrices* $A - \dfrac{1}{p}\mathrm{I}_n$ *sont inversibles. De plus, si* A *commute avec* N*, il en est de même des matrices* $A - \dfrac{1}{p}\mathrm{I}_n$*, donc d'après le point précédent :* $\det\left(\left(A - \dfrac{1}{p}\mathrm{I}_n\right) + N\right) = \det\left(A - \dfrac{1}{p}\mathrm{I}_n\right)$*. On obtient alors le résultat en faisant tendre* $p \to +\infty$ *et en utilisant la continuité du déterminant.*

À vous de jouer !

Exercice 7.1

SF7.1 **SF7.2** **SF7.3**

Soit $E = \mathscr{C}^1\big([0\,;1],\mathbb{R}\big)$. On pose, pour $f \in E$:

$$N(f) = \sqrt{f^2(0) + \int_0^1 f'^2(t)\,\mathrm{d}t} \quad \text{et} \quad N'(f) = \|f\|_\infty + \|f'\|_\infty\,.$$

Montrer que N et N' sont des normes sur E ; les comparer entre elles, ainsi qu'à la norme $\|\cdot\|_\infty$.

Exercice 7.2

SF7.1

Pour $A = (a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{K})$, on pose : $N(A) = \max\limits_{1\leqslant i\leqslant n}\left\{\sum\limits_{j=1}^{n} |a_{ij}|\right\}$.

Montrer que N est une norme d'algèbre unitaire sur $\mathcal{M}_n(\mathbb{K})$.

Exercice 7.3

SF7.5 **SF7.11** **SF7.12**

Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'un espace vectoriel normé E. Démontrer que, si $\lim\limits_{n\to\infty} \|u_n\| = +\infty$, alors l'ensemble $U = \{u_n,\, n \in \mathbb{N}\}$ est fermé.

Exercice 7.4 **SF7.24**

Soit $A \in \mathcal{M}_p(\mathbb{R})$ une matrice antisymétrique, telle que la suite $\left(A^n\right)_{n\in\mathbb{N}}$ converge vers une matrice B. Que peut-on dire de B ?

Exercice 7.5 **SF7.5**

Soit $A \in \mathcal{M}_p(\mathbb{K})$. Montrer que $\lim\limits_{n\to+\infty} \left(\mathrm{I}_n + \frac{1}{n}A\right)^n = \exp(A)$.

Exercice 7.6 **SF4.19** **SF7.7**

Soit $M \in \mathcal{M}_p(\mathbb{C})$. Démontrer que les propriétés suivantes sont équivalentes :

(i) toute valeur propre de M est de module strictement inférieur à 1 ;

(ii) $\lim\limits_{n\to+\infty} M^n = 0$;

(iii) la série $\sum\limits_{n\in\mathbb{N}} M^n$ converge.

Exercice 7.7 **SF7.8** **SF7.12**

Soit C une partie convexe d'un espace vectoriel normé E.

Montrer que $\overline{C}$ et $\overset{\circ}{C}$ sont des parties convexes.

Exercice 7.8 **SF7.11** **SF7.17**

Soit E un espace vectoriel normé de dimension finie et $(u_n)_{n\in\mathbb{N}}$ une suite bornée d'éléments de E.

Montrer que l'ensemble des valeurs d'adhérence de cette suite est une partie compacte de E.

Exercice 7.9 **SF7.12** **SF7.13** **SF7.14** **SF7.17**

Une matrice carrée $A = (a_{ij})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R})$ est dite stochastique si pour tout $(i,j) \in [\![1\,;n]\!]^2$, $a_{ij} \in [0\,;1]$ et si, pour tout $i \in [\![1\,;n]\!]$, la somme des coefficients de la i^{e} ligne de A est égale à 1.

Montrer que l'ensemble des matrices stochastiques est une partie compacte de $\mathcal{M}_n(\mathbb{R})$.

Exercice 7.10 **SF7.20**

1) Soient $(a_1,\ldots,a_p)$ des nombres complexes deux à deux distincts . Montrer que l'ensemble $\mathbb{C} \setminus \{a_1,\ldots,a_p\}$ est connexe par arcs.

2) Montrer que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs.

Exercice 7.11 **SF7.30** **SF7.31**

Soient E et F deux espaces vectoriels normés, et f une application de E dans F.

Montrer que f est uniformément continue si et seulement si, pour tout couple de suites $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ d'éléments de E telles que $\lim\limits_{n\to+\infty} \|x_n - y_n\|_E = 0$, on a $\lim\limits_{n\to+\infty} \|f(x_n) - f(y_n)\|_F = 0$.

Exercice 7.12 **SF7.13**

Soit n un entier naturel non nul, et $r \in [\![1\,;n]\!]$.

1) Montrer que l'ensemble des matrices de rang inférieur ou égal à r est une partie fermée de $\mathcal{M}_n(\mathbb{K})$.

2) Déterminer l'adhérence dans $\mathcal{M}_n(\mathbb{K})$ de l'ensemble des matrices de rang r.

Exercice 7.13 **SF7.33**

On se place dans $\mathcal{M}_p(\mathbb{K})$ muni d'une norme d'algèbre $\|\cdot\|$.

1) Montrer que pour toutes matrices A, $B \in \mathcal{M}_n(\mathbb{K})$ et tout $n \in \mathbb{N}^*$ on a :

$$\|(A+B)^n - B^n\| \leqslant n\,\|A\|\,\big(\|A\| + \|B\|\big)^{n-1}\,.$$

2) En déduire que l'application $M \mapsto \exp(M)$ est continue sur $\mathcal{M}_p(\mathbb{K})$.

Exercice 7.14 **SF7.34**

1) Pour A, $B \in \mathcal{M}_n(\mathbb{K})$, montrer que $\widetilde{AB} = \widetilde{B}\,\widetilde{A}$ où $\widetilde{M} = {}^{\mathrm{t}}\mathrm{com}(M)$ désigne la matrice complémentaire de M (commencer par traiter le cas où A et B sont inversibles).

2) Soit $n \in \mathbb{N}^*$, et soient A, B, $C \in \mathcal{M}_n(\mathbb{K})$.
Montrer que si $CD = DC$, $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - BC)$ (commencer par le cas D inversible en utilisant un produit par blocs judicieux).

Thème 8 - Suites numériques

Il s'agit ici d'un chapitre de révision. Nous nous sommes donc contentés de rappeler sur des exemples bien choisis les principales méthodes, sans vouloir être exhaustif et en détaillant surtout celles qui, par expérience, posent problème.
De plus ce chapitre permet de se familiariser avec les techniques de base de l'analyse qui seront utilisées dans toute la suite.

■ Comment étudier une suite définie sous forme explicite ?

On considère ici des suites (u_n) où chaque terme u_n est connu en fonction de n.

SF8.1 Utiliser des méthodes relatives aux fonctions numériques

Lorsque u_n est donné par une formule de la forme $u_n = f(n)$ où f est la restriction à $\mathbb{N}$ d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$, l'étude de f (variations, comportement en $+\infty$, encadrement) suffit en général pour conclure.

Exemple

Déterminer la limite de la suite u définie par : $u_n = \left(2\sin\frac{1}{n} + \frac{1}{2}\cos n\right)^n$ pour $n \in \mathbb{N}^*$.

On utilise ici l'inégalité bien connue : $\forall x \in \left[0\,;\frac{\pi}{2}\right]$, $0 \leqslant \sin x \leqslant x$ *(inégalité de convexité). On en déduit :*

$$\forall n \in \mathbb{N}^*,\ \left|2\sin\frac{1}{n} + \frac{1}{2}\cos n\right| \leqslant 2\left|\sin\frac{1}{n}\right| + \frac{1}{2}|\cos n| \leqslant \frac{2}{n} + \frac{1}{2}\cdot$$

Pour $n \geqslant 5$ *on aura donc* $|u_n| \leqslant \left(\frac{2}{5} + \frac{1}{2}\right)^n = \left(\frac{9}{10}\right)^n$ *donc* $\lim\limits_{n\to+\infty} u_n = 0$.

Exemple

Étudier le sens de variation de la suite u définie par : $\forall n \geqslant 1$, $u_n = \sqrt{n}\sin\left(\frac{1}{n}\right)$.

Pour étudier le signe de $u_{n+1} - u_n$, *on peut poser* $u_n = f(n)$ *avec* $f : x \in [1\,;+\infty[\longmapsto \sqrt{x}\sin\left(\frac{1}{x}\right)$.

f *est évidemment de classe* $\mathscr{C}^\infty$ *sur* $[1\,;+\infty[$ *et pour tout* $x \geqslant 1$:

$$f'(x) = \frac{1}{2\sqrt{x}}\sin\left(\frac{1}{x}\right) - \frac{1}{x^{3/2}}\cos\left(\frac{1}{x}\right) = \frac{1}{2\sqrt{x}}\cos\left(\frac{1}{x}\right)\left(\tan\left(\frac{1}{x}\right) - \frac{2}{x}\right),$$

donc $f'(x)$ *est du signe de* $\tan\left(\frac{1}{x}\right) - \frac{2}{x} = \varphi\left(\frac{1}{x}\right)$ *avec* $\varphi : u \mapsto \tan u - 2u$ *pour* $u \in [0\,;1]$.
L'étude des variations de φ *sur* $[0\,;1]$ *ne présente pas de difficulté, et permet d'obtenir que* φ *reste négative sur cet intervalle.*
On en déduit que $f'(x) \leqslant 0$ *pour* $x \geqslant 1$, *donc* f *est décroissante et la suite* $(u_n)_{n\in\mathbb{N}^*}$ *est aussi décroissante.*

✓ *Pour montrer que la suite u décroît à partir d'un certain rang seulement, ce qui est souvent suffisant, il était plus rapide de faire un développement limité : on trouve $u_{n+1}-u_n \underset{n\to+\infty}{\sim} -\dfrac{1}{2n^{3/2}}$, donc $u_{n+1}-u_n \leqslant 0$ pour n assez grand.*

SF8.2 Utiliser la monotonie de la suite

Commençons par une mise en garde importante.

L'étude de la monotonie d'une suite ne doit pas être systématique !

Trop d'entre vous ont en effet le mauvais réflexe de commencer l'étude d'une suite par ce point. Il faut en fait que la situation s'y prête, c'est-à-dire que la détermination du signe de $u_{n+1}-u_n$, ou l'évaluation de $\dfrac{u_{n+1}}{u_n}$ (*lorsque tous les termes sont strictement positifs ou strictement négatifs* !), soit aisée.

De plus, la détermination de la monotonie d'une suite n'a d'intérêt que si l'on sait ensuite la majorer ou la minorer (cela peut se faire assez souvent par récurrence), ou montrer qu'elle n'est pas bornée (cela peut se faire parfois par l'absurde).

Ces remarques étant faites, le théorème de la limite monotone reste cependant un outil très important.

Exemple

Pour $n \geqslant 1$, soit $u_n = \displaystyle\sum_{k=1}^{n} \frac{1}{\sqrt{k}} - 2\sqrt{n}$. Montrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ converge.

- *On calcule : $u_{n+1}-u_n = \dfrac{1}{\sqrt{n+1}} - 2(\sqrt{n+1}-\sqrt{n})$.*
  Or $\sqrt{n+1}-\sqrt{n} = \dfrac{1}{\sqrt{n+1}+\sqrt{n}} \geqslant \dfrac{1}{2\sqrt{n+1}}$, donc $u_{n+1}-u_n \leqslant 0$: la suite est décroissante.
- *De même : $\forall k \in \mathbb{N}^*$, $2\left(\sqrt{k+1}-\sqrt{k}\right) = \dfrac{2}{\sqrt{k+1}+\sqrt{k}} \leqslant \dfrac{1}{\sqrt{k}}$, donc, en additionnant les inégalités :*
  $$\sum_{k=1}^{n} \frac{1}{\sqrt{k}} \geqslant 2\sum_{k=1}^{n}\left(\sqrt{k+1}-\sqrt{k}\right) = 2\left(\sqrt{n+1}-1\right)$$
  après télescopage.
  On en déduit $u_n \geqslant 2\left(\sqrt{n+1}-1\right) - 2\sqrt{n} \geqslant -2$: la suite $(u_n)_{n\in\mathbb{N}^*}$ est minorée, et par le théorème de la limite monotone, elle converge.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ une suite bornée. Montrer que la suite de terme général :

$$v_n = \sup\{u_p \mid p \geqslant n\}$$

est convergente.

Pour tout $n \in \mathbb{N}$, notons $A_n = \{u_p \mid p \geqslant n\}$. La suite u étant bornée, les ensembles A_n le sont aussi, d'où l'existence de la borne supérieure v_n.

L'inclusion $A_{n+1} \subset A_n$ montre alors que $v_{n+1} \leqslant v_n$. La suite $(v_n)_{n\in\mathbb{N}}$ est donc décroissante ; de plus par définition on a $v_n \geqslant u_n$ donc la suite v est minorée par la borne inférieure de la suite u.

D'après le théorème de la limite monotone, elle est convergente.

✓ *La limite de la suite (v_n) s'appelle la limite supérieure de la suite (u_n) ; c'est sa plus grande valeur d'adhérence.*

SF8.3 Utiliser un encadrement

Il s'agit ici d'utiliser le théorème d'encadrement rappelé en **[S8.3]**. Si le principe est simple, obtenir des encadrements convenables n'est pas toujours facile.

Exemple

Montrer que la suite de terme général $u_n = \displaystyle\sum_{k=0}^{n} \frac{1}{\binom{n}{k}}$ converge vers une certaine limite ℓ, puis donner un équivalent de $u_n - \ell$ lorsque $n \to +\infty$.

- *On a déjà :* $u_n \geqslant \dfrac{1}{\binom{n}{0}} + \dfrac{1}{\binom{n}{n}} = 2$ *pour tout* $n \geqslant 1$.
- *D'autre part, pour* $n \geqslant 4$ *et* $k \in [\![2\,;n-2]\!]$,

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n(n-1)}{2} \prod_{i=2}^{k-1} \underbrace{\frac{n-i}{k+2-i}}_{\geqslant 1} \geqslant \frac{n(n-1)}{2},$$

donc pour $n \geqslant 4$ *:*

$$u_n \leqslant \frac{1}{\binom{n}{0}} + \frac{1}{\binom{n}{1}} + \sum_{k=2}^{n-2} \frac{2}{n(n-1)} + \frac{1}{\binom{n}{n-1}} + \frac{1}{\binom{n}{n}} \leqslant 2 + \frac{2}{n} + \frac{2(n-3)}{n(n-1)}\cdot$$

- *L'encadrement* $2 \leqslant u_n \leqslant 2 + \dfrac{2}{n} + \dfrac{2(n-3)}{n(n-1)}$ *permet alors de conclure* $\lim\limits_{n\to+\infty} u_n = 2$.
- *Pour obtenir un équivalent de* $u_n - 2$ *il suffit d'utiliser la même méthode avec un encadrement plus précis. En isolant cette fois-ci les trois premiers et trois derniers termes de la somme, on a, pour* $n \geqslant 6$ *:*

$$\begin{aligned} u_n &= \frac{1}{\binom{n}{0}} + \frac{1}{\binom{n}{1}} + \frac{1}{\binom{n}{2}} + \sum_{k=3}^{n-3} \frac{1}{\binom{n}{k}} + \frac{1}{\binom{n}{n-2}} + \frac{1}{\binom{n}{n-1}} + \frac{1}{\binom{n}{n}} \\ &= 2 + \frac{2}{n} + \frac{4}{n(n-1)} + \sum_{k=3}^{n-3} \frac{1}{\binom{n}{k}}\cdot \end{aligned}$$

Par un calcul semblable à celui fait auparavant on montre que :

$$\forall\, n \geqslant 6,\ \forall\, k \in [\![3\,;n-3]\!],\ \binom{n}{k} \geqslant \frac{n(n-1)(n-2)}{6},$$

ce qui permet d'obtenir, pour $n \geqslant 6$ *:*

$$2 + \frac{2}{n} \leqslant u_n \leqslant 2 + \frac{2}{n} + \frac{4}{n(n-1)} + \frac{6(n-5)}{n(n-1)(n-2)},$$

et d'en déduire : $\lim\limits_{n\to+\infty} \dfrac{u_n - 2}{2/n} = 1$ *soit :* $u_n - 2 \underset{n\to+\infty}{\sim} \dfrac{2}{n}\cdot$

SF8.4 Utiliser le lien suites/séries télescopiques

Cette méthode consiste à utiliser le résultat rappelé en **[S11.2]** :

la *série* $\sum_{n\in\mathbb{N}}(u_{n+1}-u_n)$ converge si et seulement si la *suite* $(u_n)_{n\in\mathbb{N}}$ converge.

L'intérêt de cette méthode est que l'on dispose d'outils supplémentaires pour étudier les séries. Cette méthode apparaît naturellement si, en étudiant la monotonie de la suite, on a trouvé une expression de $u_{n+1}-u_n$ dont le signe est compliqué à étudier mais dont on peut avoir facilement un équivalent.

Exemple

Pour $n\geqslant 1$, soit $u_n=\sum_{k=1}^{n}\frac{1}{\sqrt{k}}-2\sqrt{n}$. Montrer que la suite $(u_n)_{n\in\mathbb{N}^*}$ converge.

Pour $n\geqslant 2$, posons $v_n=u_n-u_{n-1}$. Alors $v_n=\frac{1}{\sqrt{n}}-2\sqrt{n}+2\sqrt{n-1}$ donc

$$\begin{aligned}v_n&=\frac{1}{\sqrt{n}}-2\sqrt{n}\left(1-\left(1-\frac{1}{n}\right)^{1/2}\right)\\&=\frac{1}{\sqrt{n}}-2\sqrt{n}\left(\frac{1}{2n}+O\left(\frac{1}{n^2}\right)\right)=O\left(\frac{1}{n^{3/2}}\right).\end{aligned}$$

Par comparaison à une série de Riemann, la série $\sum v_n$ est absolument convergente, donc convergente, et par conséquent la suite (u_n) *converge.*

✓ Nous avons déjà étudié cette suite en **[SF8.2]** page 399. La méthode utilisée ci-dessus montre que l'étude de la monotonie n'est pas toujours la meileure méthode pour étudier la convergence d'une suite.

SF8.5 Faire le lien suite/intégrale

Lorsque u_n se présente sous la forme $u_n=\sum_{k=1}^{n}f(k)$ avec f monotone, on peut penser à utiliser la méthode de comparaison série/intégrale pour obtenir un encadrement de u_n. Nous n'en donnerons pas ici d'exemples, vous en trouverez en **[SF11.6]** et en **[SF12.23]**, et nous l'utiliserons dans de nombreux exercices. On pouvait aussi l'utiliser pour l'exemple précédent (essayez !).

On peut aussi parfois essayer de reconnaître directement une somme de Riemann (**[S10.8]**), ou s'y ramener.

Exemple

Déterminer la limite de la suite u définie pour $n\geqslant 1$ par :

$$u_n=\sum_{k=1}^{n}\frac{1}{\sqrt{(n+k-1)(n+k)}}.$$

Pour reconnaître une somme de Riemann, on fait d'abord apparaître $\frac{1}{n}$:

$$u_n=\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\sqrt{\left(1+\frac{k-1}{n}\right)\left(1+\frac{k}{n}\right)}}.$$

Sous cette forme ce n'est pas tout à fait une somme de Riemann mais on peut s'y ramener en utilisant l'encadrement :

$$\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\sqrt{\left(1+\frac{k}{n}\right)^2}} \leqslant u_n \leqslant \frac{1}{n}\sum_{k=1}^{n}\frac{1}{\sqrt{\left(1+\frac{k-1}{n}\right)^2}}\cdot$$

Les deux termes encadrant u_n sont deux sommes de Riemann associées à la fonction continue $x \mapsto \dfrac{1}{\sqrt{(1+x)^2}} = \dfrac{1}{1+x}$ et à une subdivision régulière de l'intervalle $[0\,;1]$ en n intervalles. Elles tendent donc toutes deux vers $\displaystyle\int_0^1 \frac{\mathrm{d}x}{1+x} = \ln 2$ et, par le théorème d'encadrement, $\lim\limits_{n\to+\infty} u_n = \ln 2$.

■ **Comment étudier une suite définie par une relation de récurrence $u_{n+1} = f(u_n)$?**

SF8.6 Suivre la méthode générale

Nous exposons ici une méthode générale pour l'étude de telles suites. Dans beaucoup de cas, certaines des étapes décrites ci-dessous pourront être traitées de manière plus simple. Par exemple, lors de l'étude de la suite $u_{n+1} = u_n + u_n^2$, nul besoin d'étudier la fonction $f\colon x \mapsto x+x^2$ pour montrer que la suite est croissante !

1. **Généralités**
   Il faut que cette définition de la suite ait un sens. Pour cela, si f est définie sur une partie D de $\mathbb{R}$, on doit avoir $u_0 \in D$ et, pour tout $n \in \mathbb{N}$, $u_n \in D$. Pour cela, il *suffit* que $f(D) \subset D$, c'est-à-dire que D *soit stable par* f.
   On fera donc cette hypothèse dans tout ce qui suit.

2. **Limite éventuelle**
   Si (u_n) converge vers $\ell \in D$, et si f est continue en ℓ, alors $f(\ell) = \ell$.

   ✓ Cette propriété doit être utilisée dans presque tous les cas !

3. **Cas où f est croissante sur D (D stable par f)**
   Dans ce cas, la suite (u_n) est *monotone*.
   En effet, pour $n \geqslant 1$, $u_{n+1} - u_n = f(u_n) - f(u_{n-1})$ donc $u_{n+1} - u_n$ est du même signe que $u_n - u_{n-1}$, donc, par récurrence, du signe de $u_1 - u_0$.
   Pour étudier son sens de variation, il suffit donc de comparer u_0 et u_1. Pour cela, on peut être amené à étudier la fonction $g\colon x \mapsto f(x) - x$.
   Si f est continue, la limite éventuelle de la suite est solution de l'équation $g(x) = 0$. Si cette équation n'a pas de solution, la suite diverge ; si elle a une solution ℓ (elle peut en avoir plusieurs !), alors, par une récurrence simple :
   - $u_0 \leqslant \ell \;\Rightarrow\; u_n \leqslant \ell$ pour tout n, et la suite est majorée ;
   - $u_0 \geqslant \ell \;\Rightarrow\; u_n \geqslant \ell$ pour tout n, et la suite est minorée.

4. **Cas où f est décroissante sur D (D stable par f)**
   Dans ce cas, $f \circ f$ est croissante sur D, et D est stable par $f \circ f$. On peut donc appliquer les résultats précédents aux deux suites extraites (u_{2n}) et (u_{2n+1}).

Les deux suites extraites (u_{2n}) et (u_{2n+1}) sont alors *monotones, de sens de variations contraires* : en effet, $u_{2n+2} - u_{2n} = f(u_{2n+1}) - f(u_{2n-1})$ est de signe contraire à $u_{2n+1} - u_{2n-1}$. Pour préciser ce sens de variation, on peut être amené à étudier la fonction $g : x \mapsto f \circ f(x) - x$.

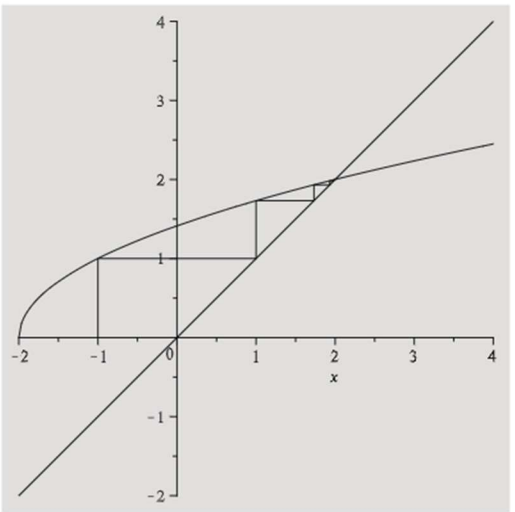
Lorsque f est continue, si les suites (u_{2n}) et (u_{2n+1}) convergent, ce ne peut être que vers une solution de l'équation $f \circ f(x) = x$. Si c'est le cas, la suite (u_n) sera convergente si et seulement si ces deux suites extraites sont adjacentes.

5. **Interprétation graphique**

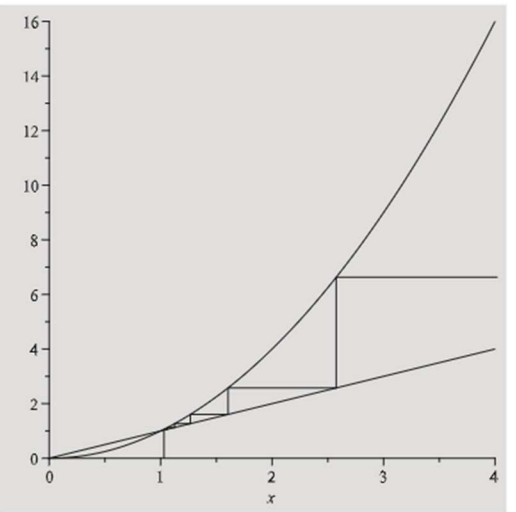
L'interprétation graphique n'est pas indispensable. Elle est surtout utile pour séparer a priori les différents cas lorsque l'on discute du comportement de la suite selon les valeurs de u_0.

Elle consiste à tracer sur une même figure la courbe représentative $\mathcal{C}$ de f et la droite $\mathcal{D}$ d'équation $y = x$; puis après avoir placé u_0 sur l'axe Ox, on obtient le point de coordonnées (u_0,u_1) sur $\mathcal{C}$; on trace un segment parallèle à Ox pour obtenir le point de coordonnées (u_1,u_1) sur $\mathcal{D}$, ce qui permet de placer u_1 sur Ox, puis on recommence pour avoir u_2 etc.

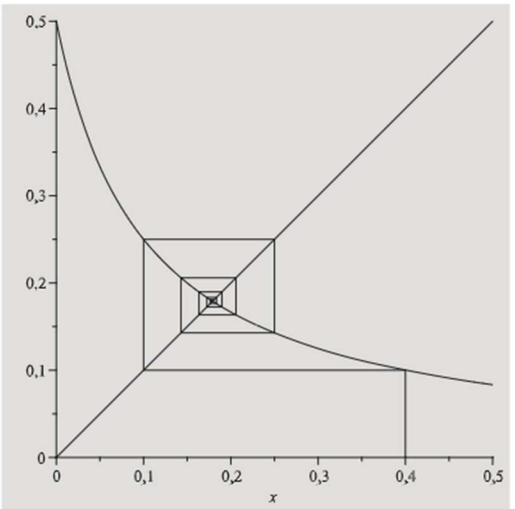
Les figures ci-après illustrent différents cas possibles lorsque f est monotone.



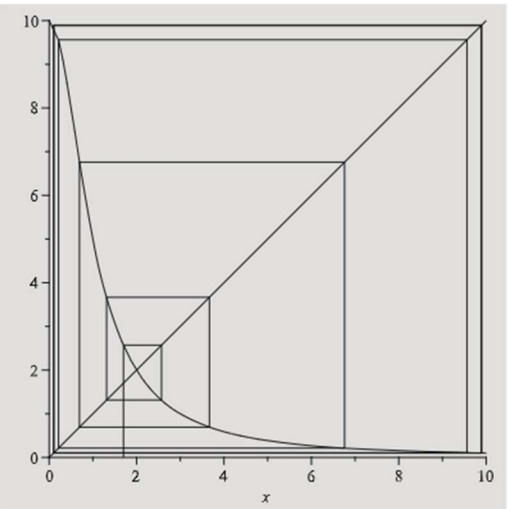
f croissante et (u_n) convergente



$f$croissante et (u_n) divergente

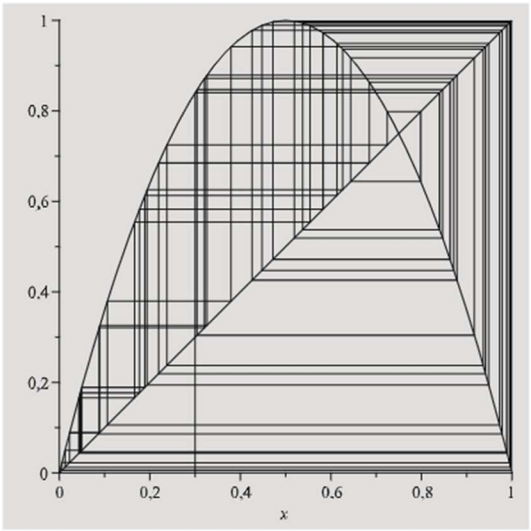


f décroissante et (u_n) convergente



f décroissante et (u_n) divergente

Lorsqu'il n'existe pas d'intervalle stable sur lequel f est monotone, il est difficile de conclure en général.



Le chaos

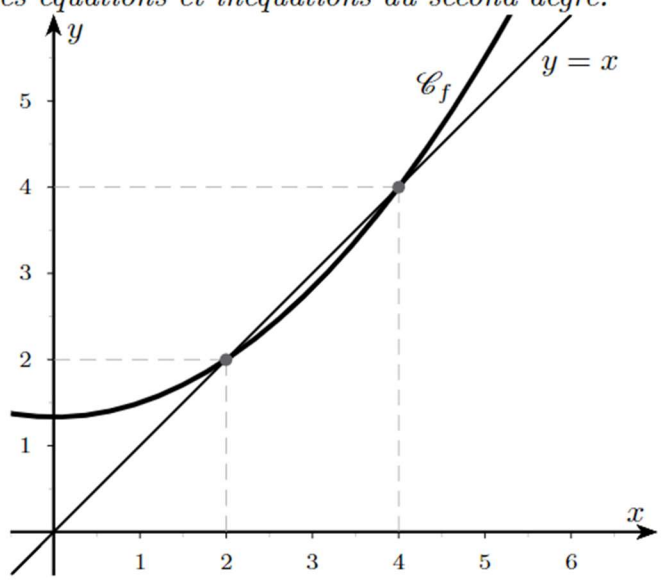
Exemple

Étude de la suite définie par : $u_0 \geqslant 0$ et $u_{n+1} = \frac{1}{6}(u_n^2 + 8)$.

La suite est définie par la relation $u_{n+1} = f(u_n)$, où $f : x \mapsto \frac{1}{6}(x^2 + 8)$.

$\mathbb{R}_+$ étant stable par f, la suite est bien définie et à valeurs dans $\mathbb{R}_+$.

Le tracé de la courbe de f est immédiat. Ce tracé permet de visualiser les deux points fixes de f ($x = 2$ et $x = 4$), ainsi que la position de $f(x)$ par rapport à x ; bien sûr, des calculs simples permettraient de démontrer rigoureusement cela, mais nous supposons que vous savez les faire : il s'agit de résoudre des équations et inéquations du second degré.



La fonction f étant croissante sur $\mathbb{R}_+$, on sait déjà que la suite $(u_n)_{n\in\mathbb{N}}$ est monotone. Son sens de variation dépend donc du signe de $u_1 - u_0 = f(u_0) - u_0$. On est donc amené à distinguer les cas suivants.

– *Si $u_0 = 2$ ou $u_0 = 4$, la suite est constante.*

– *Si $u_0 \in [0\,;2[$, puisque $f(u_0) \geqslant u_0$, la suite est croissante ; de plus, l'intervalle $[0\,;2[$ est stable par f (puisque $f([0\,;2[) = \left[\frac{4}{3}\,;2\right[$). La suite est donc majorée par 2, et elle est donc convergente. Sa limite ℓ doit vérifier : $\ell \in [0\,;2]$ et $\ell = f(\ell)$, donc $\ell = 2$.*

– *Si $u_0 \in]2\,;4[$, on trouve de la même façon que la suite est décroissante et minorée par 2 ; elle converge donc vers un réel ℓ qui appartient à $[2\,;u_0]$ (puisque la suite décroît) et qui vérifie $f(\ell) = \ell$, donc $\ell = 2$.*

– *Si $u_0 > 4$, puisque $f(u_0) > u_0$, la suite est croissante. Si elle était majorée, elle convergerait vers un réel ℓ tel que $\ell \geqslant u_0$ (puisque la suite croît) donc $\ell > 4$, et $f(\ell) = \ell$: c'est impossible. La suite étant croissante non majorée diverge donc vers $+\infty$.*

Exemple

Étude de la suite définie par : $u_0 \in \mathbb{R}_+$ et $u_{n+1} = \dfrac{10}{1+u_n^2}$.

- *La suite est définie par la relation $u_{n+1} = f(u_n)$, où $f : x \mapsto \dfrac{10}{1+x^2}$.*

  $\mathbb{R}_+$ étant stable par f, la suite est bien définie et à valeurs dans $\mathbb{R}_+$. Plus précisément, puisque $f(\mathbb{R}_+) =]0\,;10]$, les termes de la suite sont compris entre 0 et 10 à partir du rang 1 : la suite est bornée.

- *On peut chercher sa limite éventuelle en résolvant l'équation $f(x) = x$:*

  $f(x) = x \iff x^3 + x - 10 = 0 \iff (x-2)(x^2+2x+5) = 0 \iff x = 2$.

  Donc si $u_0 = 2$, la suite u est constante. On écartera désormais ce cas.

- *f étant décroissante sur $\mathbb{R}_+$, on sait que les deux suites extraites (u_{2n}) et (u_{2n+1}) sont monotones, de sens de variation contraires. Pour préciser, il nous faut étudier le signe de $f \circ f(x) - x$.*

  ✓ *Les calculs qui suivent sont un peu fastidieux, et deviennent vite compliqués si on les fait sans méthode. Nous vous invitons à les faire par vous-même, et à bien comprendre en quoi le corrigé ci-dessous minimise le nombre de calculs.*

  $$\forall x \in \mathbb{R},\ f \circ f(x) = \frac{10}{1+f(x)^2} = \frac{10}{1+\left(\frac{10}{1+x^2}\right)^2} = \frac{10(1+x^2)^2}{x^4+2x^2+101}.$$

  Donc $f \circ f(x) - x = \dfrac{10(1+x^2)^2 - x(x^4+2x^2+101)}{x^4+2x^2+101}$ *est du signe de :*

  $$\begin{aligned} 10(1+x^2)^2 - x(x^4+2x^2+101) &= -x^5+10x^4-2x^3+20x^2-101x+100 \\ &= (x^3+x-10)(-x^2+10x-1) \\ &= -(x-2)(x-\alpha)(x-\beta)(x^2+2x+5) \end{aligned}$$

  où $\alpha = 5-\sqrt{24}$ et $\beta = 5+\sqrt{24}$ sont les racines du trinôme $x^2-10x+1$.

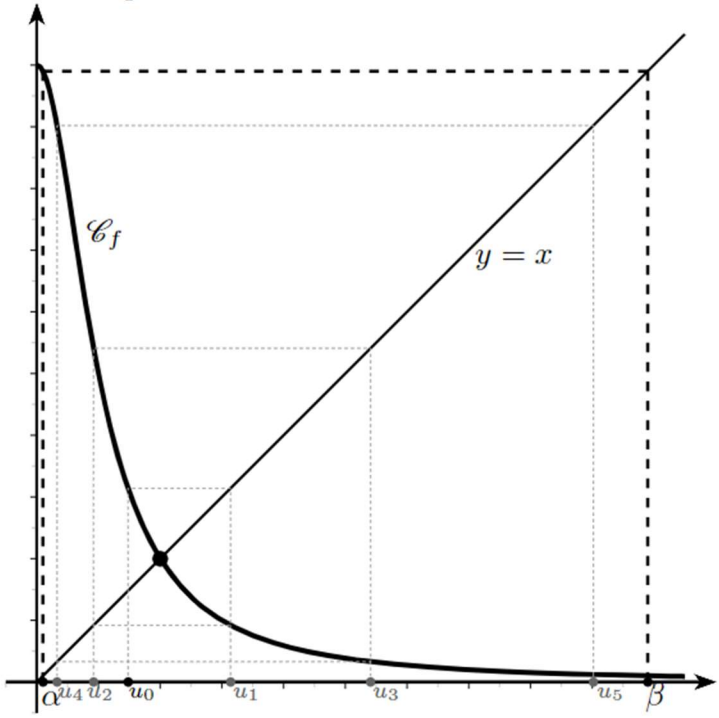
  ✓ *Pour la factorisation ci-dessus, nous avons utilisé le fait que les racines (éventuellement complexes) de l'équation $f(x) = x$ sont aussi racines*

de l'équation $f \circ f(x) = x$, donc dans l'expression de $f \circ f(x) - x$ on peut forcément factoriser $f(x) - x$.

On peut résumer le signe de $f \circ f(x) - x$ dans un tableau :

x	0		α		2		β		$+\infty$
$f \circ f(x) - x$		$+$	0	$-$	0	$+$	0	$-$	

- *Un petit graphique, même s'il n'est pas indispensable, est bien utile pour appréhender le comportement de la suite.*



- *On peut maintenant conclure, en distinguant plusieurs cas.*
  - *Si $u_0 = 2$, la suite est constante.*
  - *Si $u_0 = \alpha$, puisque $f(\alpha) = \beta$ et $f(\beta) = \alpha$, la suite (u_{2n}) est constante égale à α et la suite (u_{2n+1}) est constante égale à β.*
  - *Si $u_0 = \beta$, on obtient un résultat similaire en échangeant les rôles des deux suites d'indices pairs et impairs.*
  - *Si $u_0 \in [0\,;\alpha[$, on a $u_2 = f \circ f(u_0) > u_0$ d'après le tableau de signes précédent, donc la suite (u_{2n}) est croissante et la suite (u_{2n+1}) est décroissante. $f \circ f$ étant croissante, l'image de l'intervalle $[0\,;\alpha[$ par $f \circ f$ est l'intervalle $\big[f \circ f(0)\,;\underbrace{f \circ f(\alpha)}_{=\alpha}\big[\subset [0\,;\alpha[$. La suite (u_{2n}) reste donc dans l'intervalle $[0\,;\alpha[$ et, puisqu'elle est croissante, converge, forcément vers un point fixe de $f \circ f$ appartenant à $[0\,;\alpha]$, c'est-à-dire α. Par continuité, la suite (u_{2n+1}) converge, elle, vers $f(\alpha) = \beta$ (puisque $u_{2n+1} = f(u_{2n})$).*

    Les deux suites extraites ayant une limite différente, la suite $(u_n)_{n\in\mathbb{N}}$ est divergente.
  - *On traiterait de la même façon les cas $u_0 \in]\alpha\,;2[$, $u_0 \in]2\,;\beta[$ et $u_0 > \beta$. Vous pouvez vous entraîner, en vous aidant du graphique pour contrôler vos résultats.*

✓ L'exemple précédent peut paraître complexe, mais même s'il est long, il n'est pas trop difficile à traiter si l'on s'y prend avec méthode, soin et rigueur, en connaissant bien la méthode telle qu'elle a été exposée au début de ce paragraphe.

SF8.7 Utiliser le théorème du point fixe

Le théorème du point fixe n'est plus au programme, mais la méthode exposée ci-dessous est à retenir (il faut bien sûr refaire les démonstrations à chaque fois).

On suppose ici que f est définie sur un segment $I = [a\,;b]$, à valeurs dans I (donc I est stable par f), et que f est *contractante*, c'est-à-dire qu'il existe un réel positif $k < 1$ tel que :

$$\forall\,(x,y) \in I^2\ ,\ |f(x) - f(y)| \leqslant k\,|x - y|\ .$$

Il est important de remarquer que, lorsque f est de classe $\mathscr{C}^1$ sur I, f est contractante de rapport k si et seulement si $\sup\limits_{x\in I} |f'(x)| \leqslant k < 1$ (d'après l'inégalité des accroissements finis).

Dans ces conditions, on a les résultats suivants.

1. f admet un et un seul point fixe ℓ dans I.

   En effet : f étant contractante, elle est continue. Le théorème des valeurs intermédiaires appliqué à la fonction continue $g\colon x \mapsto f(x) - x$ assure alors l'existence de ℓ (car $g(a) \geqslant 0$ et $g(b) \leqslant 0$). Enfin, s'il existait deux points fixes ℓ et ℓ' distincts dans I, l'inégalité $|f(\ell) - f(\ell')| \leqslant k\,|\ell - \ell'|$ avec $k < 1$ conduirait à une contradiction.

2. Pour tout $u_0 \in I$, la suite définie par la relation de récurrence $u_{n+1} = f(u_n)$ converge vers ℓ.

   En effet : pour tout n on a $|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leqslant k\,|u_n - \ell|$ donc une récurrence immédiate donne $|u_n - \ell| \leqslant k^n\,|u_0 - \ell|$ pour tout n. Puisque $k \in [0\,;1[$, $\lim\limits_{n\to+\infty} k^n = 0$ d'où le résultat.

Exemple

Étudier la suite définie par : $u_0 \in \mathbb{R}$ et $u_{n+1} = \cos u_n$.

- *On commence par remarquer que, pour tout u_0, la suite u est bien définie. Pour tout $u_0 \in \mathbb{R}$, on a $u_1 \in [-1\,;1]$ puis $u_2 \in [0\,;\cos 1]$. L'intervalle $[0\,;\cos 1]$ étant stable par la fonction* cos*, on aura ensuite $u_n \in [0\,;\cos 1]$ pour tout $n \geqslant 2$.*
- *L'étude rapide de la fonction $g\colon x \mapsto \cos x - x$ montre que g est continue et strictement décroissante sur cet intervalle, et passe d'une valeur positive à une valeur négative. Elle s'annule donc une fois et une seule pour une certaine valeur ℓ de x ; ℓ est l'unique point fixe de la fonction* cos *sur cet intervalle.*
- *Sur cet intervalle, la fonction* cos *est contractante : en effet,*

$$\forall\, x \in [0\,;\cos 1],\ |\cos' x| = \sin x \leqslant \sin(\cos 1) \leqslant \sin 1 < 1\,,$$

puis d'après l'inégalité des accroissements finis :

$$\forall\,(x,y)\in[0\,;\cos 1]^2,\ |\cos x-\cos y|\leqslant(\sin 1)\,|x-y|\ ,$$

et en particulier :

$$\forall\,x\in[0\,;\cos 1],\ |\cos x-\ell|\leqslant(\sin 1)\,|x-\ell|\ .$$

On aura donc, pour $n\geqslant 2$ *:*

$$|u_{n+1}-\ell|=|\cos u_n-\ell|\leqslant(\sin 1)\,|u_n-\ell|\ ,$$

d'où par récurrence immédiate :

$$\forall\,n\geqslant 2,\ |u_n-\ell|\leqslant(\sin 1)^{n-2}\,|u_2-\ell|\ .$$

- *Puisque* $0<\sin 1<1$, $\lim\limits_{n\to+\infty}(\sin 1)^{n-2}=0$ *donc part le théorème d'encadrement, la suite* $(u_n)_{n\in\mathbb{N}}$ *converge vers* ℓ*.*

  ✓ Nous avons volontairement beaucoup détaillé les calculs dans cet exemple, car les résultats énoncés dans le cas général ne figurent pas au programme donc doivent être redémontrés. De plus, cette méthode est vraiment très performante mais souvent négligée, c'est pourquoi nous avons insisté.

SF8.8 Utiliser une série

L'utilisation de séries télescopiques, combinée avec la sommation des relations de comparaison **[S11.12]**, est une méthode efficace pour obtenir un équivalent. Nous présentons ci-dessous sur un exemple une méthode qui s'applique à de nombreuses suites définies par récurrence.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ définie par $u_0\in\left]0\,;\frac{\pi}{2}\right]$ et la relation de récurrence $u_{n+1}=\sin u_n$. Déterminer sa limite, puis un équivalent de u_n, puis un développement asymptotique à deux termes.

- *L'intervalle* $\left]0\,;\frac{\pi}{2}\right]$ *étant stable pas la fonction* sin*, la suite reste à valeurs dans* $\left]0\,;\frac{\pi}{2}\right]$ *; puisque* $\sin x<x$ *pour* $x\in\left]0\,;\frac{\pi}{2}\right]$*, la suite est (strictement) décroissante ; étant minorée par* 0*, elle converge vers un réel* $\ell\in\left[0\,;\frac{\pi}{2}\right]$ *tel que* $\sin\ell=\ell$*, c'est-à-dire* $\ell=0$*.*
- *Soit* $\alpha\in\mathbb{R}$*. Puisque* $u_n>0$ *pour tout* n*, on peut considérer* u_n^α *; on calcule alors, à l'aide d'un développement limité (licite car* $\lim\limits_{n\to+\infty}u_n=0$*) :*

$$\begin{aligned}u_{n+1}^\alpha-u_n^\alpha=(\sin u_n)^\alpha-u_n^\alpha&=\left(u_n-\frac{u_n^3}{6}+o\big(u_n^3\big)\right)^\alpha-u_n^\alpha\\&=u_n^\alpha\left(\left(1-\frac{u_n^2}{6}+o\big(u_n^2\big)\right)^\alpha-1\right)=-\alpha\frac{u_n^{\alpha+2}}{6}+o\big(u_n^{\alpha+2}\big)\ .\end{aligned}$$

  On en déduit que, si l'on choisit $\alpha=-2$*, on a* $\lim\limits_{n\to+\infty}\big(u_{n+1}^{-2}-u_n^{-2}\big)=\dfrac{1}{3}$*.*

Par le théorème de comparaison des sommes partielles de séries divergentes, on en déduit : $u_n^{-2} \underset{n\to+\infty}{\sim} \dfrac{n}{3}$ *et finalement :*

$$u_n \underset{n\to+\infty}{\sim} \sqrt{\frac{3}{n}}\,.$$

- *Pour obtenir le terme suivant du développement asymptotique, on utilise le même principe, en utilisant un développement limité à un ordre supérieur. Posons* $v_n = u_n^{-2} - \dfrac{n}{3}$*. On a alors :*

$$\begin{aligned} v_{n+1} - v_n = u_{n+1}^{-2} - u_n^{-2} - \frac{1}{3} &= (\sin u_n)^{-2} - u_n^{-2} - \frac{1}{3} \\ &= u_n^{-2}\left(1 - \frac{u_n^2}{6} + \frac{u_n^4}{120} + o\big(u_n^5\big)\right)^{-2} - u_n^{-2} - \frac{1}{3} = \frac{1}{15}u_n^2 + o\big(u_n^3\big)\,, \end{aligned}$$

en ayant utilisé : $(1+x)^{-2} \underset{0}{=} 1 - 2x + 3x^2 + o\big(x^2\big)$*.*

Puisque $u_n^2 \sim \dfrac{3}{n}$*, on a* $v_{n+1} - v_n \sim \dfrac{1}{5n}$*. La série* $\sum \dfrac{1}{5n}$ *étant divergente, le même théorème donne* $v_n \sim \displaystyle\sum_{k=1}^{n} \frac{1}{5k} \sim \frac{1}{5}\ln n$*, c'est-à-dire* $v_n = \dfrac{\ln n}{5} + o(\ln n)$ *d'où :*

$$\begin{aligned} u_n = \left(\frac{n}{3} + v_n\right)^{-1/2} &= \left(\frac{n}{3} + \frac{\ln n}{5} + o(\ln n)\right)^{-1/2} \\ &= \left(\frac{n}{3}\right)^{-1/2}\left(1 + \frac{3}{5}\frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)\right)^{-1/2} \\ &= \left(\frac{n}{3}\right)^{-1/2}\left(1 - \frac{3\ln n}{10n} + o\left(\frac{\ln n}{n}\right)\right). \end{aligned}$$

Le développement asymptotique cherché est donc :

$$u_n = \sqrt{\frac{3}{n}} - \frac{3\sqrt{3}\,\ln n}{10\,n^{3/2}} + o\left(\frac{\ln n}{n^{3/2}}\right).$$

■ Comment étudier deux suites définies simultanément ?

SF8.9 Utiliser directement les résultats sur les suites adjacentes

Pour démontrer que deux suites sont adjacentes, on peut dans l'idéal utiliser directement la définition. C'est en général le cas lorsque u_n et v_n sont directement donnés en fonction de n.

Exemple

Étudier la convergence des deux suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ données par :

$$\forall\, n \in \mathbb{N}^*,\ u_n = \sum_{k=0}^{n} \frac{1}{k!} \quad \text{et} \quad v_n = u_n + \frac{1}{n.n!}\,.$$

- *Pour* $n \geqslant 1$*,* $u_{n+1} - u_n = \dfrac{1}{(n+1)!}$*, donc la suite* u *est croissante.*

- *Pour tout entier $n \in \mathbb{N}^*$:*

$$\begin{aligned} v_{n+1} - v_n &= u_{n+1} - u_n + \frac{1}{(n+1)(n+1)!} - \frac{1}{n.n!} \\ &= \frac{1}{(n+1)!} + \frac{1}{(n+1)(n+1)!} - \frac{1}{n.n!} \\ &= \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)^2} - \frac{1}{n}\right) \\ &= \frac{1}{n!}\left(\frac{n(n+1)+n-(n+1)^2}{n(n+1)^2}\right) = -\frac{1}{n(n+1)^2 n!} \end{aligned}$$

  donc $v_{n+1} - v_n < 0$: la suite v est décroissante.
- *Clairement :* $\lim\limits_{n\to+\infty} (v_n - u_n) = 0$.
- *Les deux suites u et v sont donc adjacentes ; d'après le cours, elles convergent, vers la même limite.*

  ✓ *En utilisant le développement en série entière de la fonction* exp, *on voit que cette limite commune est égale à* e.

SF8.10 Utiliser la monotonie

Dans le cas (fréquent) où l'on ne connaît pas u_n ou v_n en fonction de n et pour montrer que les deux suites sont adjacentes, la démonstration de la propriété $\lim\limits_{n\to+\infty} (u_n - v_n) = 0$ n'est pas facile. On est donc ramené à étudier séparément les deux suites u et v, puis à montrer que leur limite est la même.
On pourra employer aussi cette méthode même lorsque les suites ne sont pas adjacentes.

Exemple

Étudier les suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par : $u_0 > 0$, $v_0 > 0$ et :

$$\forall\, n \in \mathbb{N},\ u_{n+1} = \frac{u_n}{2} + \frac{2}{v_n}\ ,\ v_{n+1} = \frac{1}{2u_n} + 2v_n\ .$$

- *Par récurrence, les deux suites u et v sont bien définies et à valeurs strictement positives.*
- *On a donc $v_{n+1} > 2v_n$ d'où par récurrence $v_n > 2^n v_0$. On en déduit* $\lim\limits_{n\to+\infty} v_n = +\infty$
- *Le calcul de $u_{n+1} - u_n$, qui sera détaillé par la suite, fait intervenir le produit $u_n v_n$. Il est donc normal de commencer par étudier cette suite.*

  En faisant le produit des deux relations données on obtient :

$$u_{n+1}v_{n+1} = \frac{17}{4} + u_n v_n + \frac{1}{u_n v_n}\cdot$$

  La suite $(u_n v_n)$ est donc croissante ; si elle était majorée, elle serait convergente vers un réel ℓ strictement positif (car $\ell \geqslant u_0 v_0$) tel que :

$$\ell = \frac{17}{4} + \ell + \frac{1}{\ell}\ ,$$

  ce qui est impossible. La suite étant croissante non majorée, on a $\lim\limits_{n\to+\infty} u_n v_n = +\infty$.

- $u_{n+1}-u_n=\dfrac{2}{v_n}-\dfrac{u_n}{2}=\dfrac{4-u_nv_n}{2v_n}$. *Puisque* $\lim\limits_{n\to+\infty}(4-u_nv_n)=-\infty$, *la différence* $u_{n+1}-u_n$ *sera négative à partir d'un certain rang, c'est-à-dire que la suite* u *est décroissante au moins à partir d'un certain rang.*

  Étant minorée par 0, *elle est convergente, vers un certain réel* ℓ. *En passant à la limite dans la relation de récurrence* $u_{n+1}=\dfrac{u_n}{2}+\dfrac{2}{v_n}$, *on obtient* $\ell=\dfrac{\ell}{2}$ *d'où* $\ell=0$.

Conclusion : $\lim\limits_{n\to+\infty}v_n=+\infty$ *et* $\lim\limits_{n\to+\infty}u_n=0$.

SF8.11 Calculer les termes généraux des deux suites

Lorsque les deux suites u et v sont reliées par un système de récurrence linéaire de la forme :

$$\begin{cases} u_{n+1}=au_n+bv_n \\ v_{n+1}=cu_n+du_n \end{cases},$$

on sait alors exprimer u_n et v_n en fonction de n ; la méthode est exposée en détails en **[SF4.20]**.

On peut aussi arriver à exprimer u_n et v_n en fonction de n dans certains cas particuliers.

Exemple

Étudier les suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par : $0<u_0<v_0$ et :

$$\forall n\in\mathbb{N},\ u_{n+1}=\frac{u_n^2}{u_n+v_n}\ ,\ v_{n+1}=\frac{v_n^2}{u_n+v_n}\ .$$

Par récurrence, les deux suites u *et* v *sont bien définies et à valeurs strictement positives.*

$\forall n\in\mathbb{N},\ u_{n+1}-v_{n+1}=\dfrac{u_n^2-v_n^2}{u_n+v_n}=u_n-v_n$, *donc la suite* $u-v$ *est constante :*

$$\forall n\in\mathbb{N},\ u_n-v_n=u_0-v_0 \quad (1)\,.$$

Puis $\dfrac{u_{n+1}}{v_{n+1}}=\left(\dfrac{u_n}{v_n}\right)^2$ *donc par récurrence :*

$$\forall n\in\mathbb{N},\ \frac{u_n}{v_n}=\left(\frac{u_0}{v_0}\right)^{2^n} \quad (2)\,.$$

Les relations (1) *et* (2) *donnent* $v_n=\dfrac{v_0-u_0}{1-\left(\dfrac{u_0}{v_0}\right)^{2^n}}$ *(le dénominateur n'est pas nul puisque* $u_0\neq v_0$*), puis* $u_n=u_0-v_0+v_n$.

Puisque $0<u_0<v_0$, $\lim\limits_{n\to+\infty}\left(\dfrac{u_0}{v_0}\right)^{2^n}=0$ *donc* $\lim\limits_{n\to+\infty}v_n=v_0-u_0$ *puis* $\lim\limits_{n\to+\infty}u_n=0$.

SF8.12 Raisonner par l'absurde pour prouver qu'une suite diverge

Dans des exemples précédents pages 404 et 410, nous avons déjà raisonné par l'absurde pour montrer qu'une suite diverge vers $+\infty$: si $(u_n)_{n\in\mathbb{N}}$ est une suite croissante, elle converge si et seulement si elle est majorée. Pour montrer qu'elle diverge, on peut donc par l'absurde supposer qu'elle est majorée (ou, directement, qu'elle converge) et aboutir à une contradiction.
Donnons un autre exemple.

Exemple

Soit α un réel n'appartenant pas à $\pi\mathbb{Z}$.

Montrer que les deux suites de terme général $u_n = \cos n\alpha$ et $v_n = \sin n\alpha$ divergent.

Nous allons raisonner par l'absurde ; mais il faut faire attention : la négation de la phrase :

« les suites u et v divergent »

est :

« la suite u ou *la suite v converge ».*

Supposons donc que l'une des deux suites converge.

- *Montrons d'abord que l'autre converge aussi.*
  Si, par exemple, la suite u converge vers un certain réel ℓ, alors, puisque $\alpha \notin \pi\mathbb{Z}$, en utilisant les formules trigonométriques d'addition on a :
  $$\sin n\alpha = \frac{\cos n\alpha \cos\alpha - \cos(n+1)\alpha}{\sin\alpha},$$
  donc $\lim\limits_{n\to+\infty} \sin n\alpha = \dfrac{\ell\cos\alpha - \ell}{\sin\alpha}$ *existe aussi.*
  On procède de la même façon pour démontrer que si la suite v converge, il en est de même de la suite u.
- *Supposons donc* $\lim\limits_{n\to+\infty} \cos n\alpha = \ell$ *et* $\lim\limits_{n\to+\infty} \sin n\alpha = \ell'$.
  En passant à la limite dans les formules :
  $$\begin{cases} \cos(n+1)\alpha = \cos n\alpha\cos\alpha - \sin n\alpha\sin\alpha \\ \sin(n+1)\alpha = \sin n\alpha\cos\alpha + \sin\alpha\cos n\alpha \end{cases}$$
  on obtient le système :
  $$\begin{cases} \ell = \ell\cos\alpha - \ell'\sin\alpha \\ \ell' = \ell'\cos\alpha + \ell\sin\alpha \end{cases} \quad \textit{soit} \quad \begin{cases} \ell(\cos\alpha - 1) - \ell'\sin\alpha = 0 \\ \ell\sin\alpha + \ell'(\cos\alpha - 1) = 0 \end{cases}.$$
  Ce système homogène, de déterminant $(\cos\alpha - 1)^2 + \sin^2\alpha = 2(1-\cos\alpha)$ *non nul admet la seule solution* $\ell = \ell' = 0$.
  Mais la relation $\cos^2 n\alpha + \sin^2 n\alpha = 1$ *implique par passage à la limite :* $\ell^2 + \ell'^2 = 1$, *d'où la contradiction cherchée.*

SF8.13 Utiliser des suites extraites

On sait que si une suite converge, toutes les suites extraites convergent, vers la même limite. Si donc on peut extraire d'une suite $(u_n)_{n\in\mathbb{N}}$ deux suites ayant des limites différentes, la suite u sera divergente. C'est par cette méthode que l'on peut montrer, par exemple, que la suite $n \mapsto (-1)^n$ diverge.
Donnons un autre exemple d'utilisation de ce résultat.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ définie par :

$$u_0 \in]0\,;2[\quad \text{et} \quad u_{n+1} = \sqrt{2 + (-1)^{n+1} u_n} \quad \text{pour tout } n \in \mathbb{N}\,.$$

Montrer que cette suite diverge.

- *Montrons déjà par récurrence sur n la propriété :*

  « pour tout $n \in \mathbb{N}$, u_n existe et $u_n \in]0\,;2[$ ».

  – *Cette propriété est vérifiée pour $n = 0$.*

  – *Si elle est vérifiée au rang n alors :*

  $$0 < 2 - u_n \leqslant 2 + (-1)^{n+1} u_n \leqslant 2 + u_n < 4\,,$$

  ce qui implique que la propriété sera encore vraie au rang $n+1$.

- *Si par l'absurde $(u_n)_{n\in\mathbb{N}}$ était convergente vers un réel ℓ, il en serait de même de ses suites extraites (u_{2n}) et (u_{2n+1}). Or :*

  $$u_{2n+1} = \sqrt{2 + u_{2n}} \quad \text{et} \quad u_{2n} = \sqrt{2 - u_{2n-1}}$$

  donc en passant à la limite dans ces relations on obtient :

  $$\ell = \sqrt{2 + \ell} \quad \text{et} \quad \ell = \sqrt{2 - \ell}\,,$$

  ce qui est impossible. On en déduit que la suite $(u_n)_{n\in\mathbb{N}}$ diverge.

SF8.14 Revenir à la définition avec des ε

Utiliser la définition de la limite d'une suite est toujours un peu délicat. Cela doit être fait avec rigueur et à bon escient.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à termes dans $\mathbb{Z}$.

Montrer que (u_n) converge si et seulement si (u_n) est stationnaire.

- *Évidemment, si la suite $(u_n)_{n\in\mathbb{N}}$ est stationnaire, elle converge.*
- *Réciproquement, supposons $(u_n)_{n\in\mathbb{N}}$ à valeurs entières et convergente, et soit $\ell = \lim\limits_{n\to+\infty} u_n$. Par définition de la limite (avec $\varepsilon = \frac{1}{2}$), on a :*

  $$\exists\, n_0 \in \mathbb{N} \text{ tel que } \forall\, n \geqslant n_0,\ |u_n - \ell| < \frac{1}{2}\,.$$

  Comme u_n est entier et qu'il ne peut y avoir qu'un entier strictement compris entre $\ell - \frac{1}{2}$ et $\ell + \frac{1}{2}$, u_n est égal à cet entier pour $n \geqslant n_0$ c'est-à-dire que $(u_n)_{n\in\mathbb{N}}$ est stationnaire.

Exemple

Soit $(u_n)_{n\in\mathbb{N}}$ une suite telle que les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent vers la même limite ℓ.

Montrer que la suite u converge aussi vers ℓ.

Soit $\varepsilon > 0$. D'après les hypothèses et par définition de la limite :

$$\begin{aligned} &\exists n_0 \in \mathbb{N} \quad \textit{tel que} \quad \forall n \geqslant n_0,\ |u_{2n} - \ell| < \varepsilon \\ \textit{et}\quad &\exists n_1 \in \mathbb{N} \quad \textit{tel que} \quad \forall n \geqslant n_1,\ |u_{2n+1} - \ell| < \varepsilon\,. \end{aligned}$$

En prenant $N = \max(2n_0, 2n_1 + 1)$ on aura donc :

$$\forall n \geqslant N,\ |u_n - \ell| < \varepsilon\,,$$

ce qui signifie par définition que la suite u converge vers ℓ.

✓ Ce résultat est classique et peut être considéré comme faisant partie du cours (mais il faut savoir le redémontrer si on vous le demande).

Exemple

Soit $(u_n)_{n\in\mathbb{N}^*}$ une suite de nombres complexes convergeant vers $\ell \in \mathbb{C}$, et $(a_n)_{n\in\mathbb{N}^*}$ une suite de réels strictement positifs telle que la série $\sum_{n\geqslant 1} a_n$ diverge.

Montrer que la suite (v_n) définie par : $v_n = \dfrac{\sum_{k=1}^{n} a_k u_k}{\sum_{k=1}^{n} a_k}$ converge aussi vers ℓ (ce résultat porte le nom de *théorème de Cesàro*).

Soit ε un réel strictement positif fixé. Par hypothèse :

$$\exists n_0 \in \mathbb{N} \quad \textit{tel que} \quad \forall n \in \mathbb{N},\ n \geqslant n_0 \implies |u_n - \ell| < \frac{\varepsilon}{2}\cdot$$

Pour $n \geqslant n_0$, on a :

$$v_n - \ell = \frac{\sum_{k=1}^{n} a_k u_k}{\sum_{k=1}^{n} a_k} - \ell = \frac{\sum_{k=1}^{n} a_k (u_k - \ell)}{\sum_{k=1}^{n} a_k} = \frac{\sum_{k=1}^{n_0-1} a_k (u_k - \ell) + \sum_{k=n_0}^{n} a_k (u_k - \ell)}{\sum_{k=1}^{n} a_k}\,,$$

d'où, en utilisant l'inégalité triangulaire et le fait que les a_k sont positifs :

$$\begin{aligned} |v_n - \ell| &\leqslant \frac{\sum_{k=1}^{n_0-1} a_k |u_k - \ell|}{\sum_{k=1}^{n} a_k} + \frac{\sum_{k=n_0}^{n} a_k |u_k - \ell|}{\sum_{k=1}^{n} a_k} \\ &\leqslant \frac{\sum_{k=1}^{n_0-1} a_k |u_k - \ell|}{\sum_{k=1}^{n} a_k} + \frac{\sum_{k=n_0}^{n} a_k \frac{\varepsilon}{2}}{\sum_{k=1}^{n} a_k} \leqslant \frac{\sum_{k=1}^{n_0-1} a_k |u_k - \ell|}{\sum_{k=1}^{n} a_k} + \frac{\varepsilon}{2}\cdot \end{aligned}$$

Or le premier terme de cette dernière somme tend vers 0 quand $n \to +\infty$, puisque le numérateur $\sum_{k=1}^{n_0-1} a_k |u_k - \ell|$ ne dépend pas de n et que le dénominateur $\sum_{k=1}^{n} a_k$ tend vers $+\infty$ par hypothèse (car la série à termes positifs $\sum a_n$ diverge). Il existe donc un entier n_1 (que l'on peut supposer $\geqslant n_0$) tel que, pour $n \geqslant n_1$, ce terme soit inférieur à $\frac{\varepsilon}{2}$.
Finalement, pour $n \geqslant n_1$ on aura $|v_n - \ell| < \varepsilon$, ce qui donne le résultat annoncé.

SF8.15 Utiliser le théorème de Bolzano-Weierstrass

Il faut penser à utiliser ce théorème lorsqu'une suite bornée apparaît dans un problème.

Exemple

Soit x un nombre irrationnel. On suppose que x est limite d'une suite de rationnels : $x = \lim\limits_{n\to+\infty} \dfrac{p_n}{q_n}$ avec $p_n \in \mathbb{Z}$, $q_n \in \mathbb{N}^*$.

Montrer que $\lim\limits_{n\to+\infty} q_n = +\infty$ et $\lim\limits_{n\to+\infty} |p_n| = +\infty$.

Posons $u_n = \dfrac{p_n}{q_n}$.

- *Montrons d'abord que $\lim\limits_{n\to+\infty} q_n = +\infty$, en raisonnant par l'absurde. En écrivant la négation de la définition de la limite, on obtient :*

$$\exists A \in \mathbb{R},\ \forall N \in \mathbb{N},\ \exists n \geqslant N \text{ tel que } q_n \leqslant A\,.$$

En appliquant cette propriété à $N = 0$ on obtient un entier n_0 tel que $q_{n_0} \leqslant A$; puis en l'appliquant avec $N = n_0 + 1$, on obtient un entier $n_1 > n_0$ tel que $q_{n_1} \leqslant A$ etc...
On construit ainsi une suite extraite de $(q_n)_{n\in\mathbb{N}}$ qui est bornée ; d'après le théorème de Bolzano-Weierstrass, on peut donc de nouveau extraire de cette suite une suite $q_{\varphi(n)}$ qui est convergente.
Mais comme il s'agit d'une suite d'entiers, elle est stationnaire (cf. le 1er exemple de **[SF8.14]** *page 413). Puisque $p_{\varphi(n)} = q_{\varphi(n)}u_{\varphi(n)}$, la suite $(p_{\varphi(n)})$ est convergente donc, comme il s'agit d'une suite d'entiers, elle est stationnaire.*
On peut donc extraire de la suite (u_n) une suite constante, ce qui implique que sa limite x est égale à cette constante donc est un nombre rationnel : contradiction.

- *On a donc prouvé que $\lim\limits_{n\to+\infty} q_n = +\infty$. Puisque $p_n = u_n q_n$ et que (u_n) converge vers x, on en déduit que $\lim\limits_{n\to+\infty} |p_n| = +\infty$.*

À vous de jouer !

Exercice 8.1 — SF8.7

On pose : $u_0 > 0$ et $\forall n \in \mathbb{N}$, $u_{n+1} = 1 + \frac{1}{4}\sin\frac{1}{u_n}$. Montrer que la suite u converge.

Exercice 8.2 — SF8.6

On considère la fonction numérique f de variable réelle définie par :

$$\forall x \in \mathbb{R},\ f(x) = -x^2 + 2x + 1\,,$$

et la suite $(u_n)_{n\in\mathbb{N}}$ définie par :

$$u_0 \in \mathbb{R} \quad \text{et} \quad \forall n \in \mathbb{N},\ u_{n+1} = f(u_n)\,.$$

1) Étudier rapidement la fonction f, tracer sa courbe représentative (C) et préciser les points d'intersection de (C) avec la droite d'équation $y = x$.
Par la suite, on notera ℓ_1 et ℓ_2 les abscisses de ces points ($\ell_1 < \ell_2$).

2) Étudier la suite dans le cas : $u_0 < \ell_1$.

3) Étudier la suite dans le cas : $u_0 \in]1,2[$.

Exercice 8.3 — SF8.2

Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle telle que : $\forall n \in \mathbb{N}$, $\frac{1}{2} < x_n < 1$.
On définit une suite $(y_n)_{n\in\mathbb{N}}$ par :

$$y_0 = x_0 \quad \text{et} \quad \forall n \in \mathbb{N}^*,\ y_n = \frac{x_n + y_{n-1}}{1 + x_n y_{n-1}}\,.$$

Montrer que la suite $(y_n)_{n\in\mathbb{N}}$ converge et préciser sa limite.

Exercice 8.4 — SF8.10

Étudier les suites réelles $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par : $u_0 > 0$, $v_0 > 0$ et :

$$\forall n \in \mathbb{N},\ u_{n+1} = \frac{u_n + v_n}{2}\,,\ v_{n+1} = \frac{2u_n v_n}{u_n + v_n}\,.$$

Exercice 8.5 — SF8.3 SF8.10

1) Soient x, y, $z \geqslant 0$. Montrer que $x^3 + y^3 + z^3 - 3xyz \geqslant 0$ (mettre $x + y + z$ en facteur).

2) Étudier les suites réelles (u_n), (v_n) et (w_n) définies par :

$$0 < u_0 < v_0 < w_0$$

et les relations de récurrence :

$$\forall n \in \mathbb{N},\ \frac{3}{u_{n+1}} = \frac{1}{u_n} + \frac{1}{v_n} + \frac{1}{w_n}\,,\ v_{n+1} = \sqrt[3]{u_n v_n w_n}\,,\ w_{n+1} = \frac{u_n + v_n + w_n}{3}\,.$$

Exercice 8.6 SF8.13 SF8.14

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles strictement positives.

1) Montrer que, si $\dfrac{u_{n+1}}{u_n}$ possède une limite $\ell \in \overline{\mathbb{R}}$, il en est de même de la suite $(\sqrt[n]{u_n})$.

2) Montrer que la réciproque de la propriété précédente est fausse en considérant la suite (u_n) définie par :

$$\forall\, n \in \mathbb{N},\ u_{2n} = a^n b^n \quad \text{et} \quad u_{2n+1} = a^n b^{n+1}\ (0 < a < b)\,.$$

Exercice 8.7 SF8.2 SF8.3

Pour $x \in [0\,;1]$ et n entier $\geqslant 2$, on pose $f_n(x) = x^n - nx + 1$.

1) Montrer qu'il existe un et un seul $x_n \in [0\,;1]$ tel que $f_n(x_n) = 0$.

2) **a)** Montrer que la suite $(x_n)_{n\in\mathbb{N}}$ converge.

b) Soit $a \in\ ]0\,;1[$. Étudier le signe de $f_n(a)$ pour n suffisamment grand. En déduire la limite de la suite (x_n).

3) **a)** Donner un équivalent de x_n quand n tend vers $+\infty$.

b) Donner un développement asymptotique à deux termes de x_n quand n tend vers $+\infty$.

Exercice 8.8 SF8.8 SF8.12

On considère la suite réelle $(u_n)_{n\in\mathbb{N}}$ définie par :

$$u_0 > 0 \quad \text{et} \quad \forall\, n \in \mathbb{N},\ u_{n+1} = u_n + \frac{1}{u_n}\,.$$

1) Montrer que $\lim\limits_{n\to+\infty} u_n = +\infty$.

2) Donner un développemtn asymptotique à deux termes de u_n quand n tend vers $+\infty$.

Exercice 8.9 SF8.14

Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle et $k \in\ ]0\,;1[$. On pose :

$$y_0 = x_0 \quad \text{et} \quad \forall\, n \geqslant 1,\ y_n = x_n - kx_{n-1}\,.$$

Montrer que : $\lim\limits_{n\to+\infty} x_n = 0 \iff \lim\limits_{n\to+\infty} y_n = 0$ (s'inspirer de la démonstration du théorème de Cesàro).

Exercice 8.10 SF8.13 SF8.15

Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle bornée telle que $\lim\limits_{n\to+\infty} \left(x_n + \dfrac{x_{2n}}{2}\right) = 1$.

Montrer que : $\lim\limits_{n\to+\infty} x_n = \dfrac{2}{3}$.

Thème 9 - Fonctions réelles d'une variable réelle

Le début de ce chapitre est consacré à des révisions sur les fonctions réelles d'une variable réelle. Ce thème ayant été étudié en 1re année, nous nous sommes contentés de rappeler certains points parmi ceux qui nous semblent les plus importants.

■ Comment étudier et utiliser la dérivabilité d'une fonction ?

SF9.1 Revenir à la définition

On peut bien sûr revenir à la définition de la dérivabilité en un point, avec la limite du taux d'accroissement, pour étudier la dérivabilité d'une fonction donnée en un point qui pose problème.
Mais on peut aussi utiliser cette définition dans des exercices plus théoriques ; dans ce cas, il est souvent commode d'exprimer la dérivabilité à l'aide d'un développement limité d'ordre 1.

Exemple

Soit $f :]-1\,;1[\to \mathbb{R}$, dérivable en 0, et telle que $f(0) = 0$.

1) Pour tout $n \in \mathbb{N}^*$, on pose $u_n = \sum_{k=1}^{n} f\left(\frac{k}{n^2}\right)$. Calculer $\lim\limits_{n\to+\infty} u_n$.

2) Donner un équivalent quand n tend vers $+\infty$ de $v_n = \prod_{k=1}^{n} (n^2 + k)$.

1) *Par définition, dire que f est dérivable en 0 peut s'écrire :*

$$f(x) = f(0) + xf'(0) + x\varphi(x)\,,$$

avec $\lim\limits_{x\to 0} \varphi(x) = 0$. Ici, $f0) = 0$, on a donc :

$$u_n = \sum_{k=1}^{n} \frac{k}{n^2} f'(0) + \sum_{k=1}^{n} \frac{k}{n^2} \varphi\left(\frac{k}{n^2}\right).$$

D'une part :

$$\sum_{k=1}^{n} \frac{k}{n^2} f'(0) = \frac{f'(0)}{n^2} \sum_{k=1}^{n} k = \frac{f'(0)}{n^2} \cdot \frac{n(n+1)}{2} \underset{n\to+\infty}{\longrightarrow} \frac{f'(0)}{2}\,.$$

D'autre part, soit $\varepsilon > 0$; φ ayant une limite nulle en 0, il existe un voisinage V de 0 tel que, pour tout $x \in V$, $|\varphi(x)| < \varepsilon$. Or il existe un entier N tel que $\left[-\frac{1}{N}\,;\frac{1}{N}\right] \subset V$, donc pour $n \geqslant N$ on a :

$$\left|\sum_{k=1}^{n} \frac{k}{n^2} \varphi\left(\frac{k}{n^2}\right)\right| \leqslant \sum_{k=1}^{n} \frac{k}{n^2} \left|\varphi\left(\frac{k}{n^2}\right)\right| \leqslant \varepsilon \sum_{k=1}^{n} \frac{k}{n^2} \leqslant \epsilon$$

(en effet, pour $1 \leqslant k \leqslant n$, $\frac{k}{n^2} \in \left[0\,;\frac{1}{n}\right] \subset \left[-\frac{1}{N}\,;\frac{1}{N}\right]$).

Cela signifie que $\lim\limits_{n\to+\infty}\left(\sum\limits_{k=1}^{n}\frac{k}{n^2}\varphi\left(\frac{k}{n^2}\right)\right)=0$. *On en déduit que* $\lim\limits_{n\to+\infty}u_n=\frac{f'(0)}{2}$.

2) *Pour* $n\in\mathbb{N}^*$, *on a* $v_n=\prod\limits_{k=1}^{n}n^2\left(1+\frac{k}{n^2}\right)=n^{2n}\prod\limits_{k=1}^{n}\left(1+\frac{k}{n^2}\right)$, *donc* $\ln\left(\frac{v_n}{n^{2n}}\right)=\sum\limits_{k=1}^{n}\ln\left(1+\frac{k}{n^2}\right)$.

En appliquant le résultat de la question précédente à la fonction $x\mapsto\ln(1+x)$, *on obtient :*

$$\lim_{n\to+\infty}\ln\left(\frac{v_n}{n^{2n}}\right)=\frac{1}{2}\text{, donc }\lim_{n\to+\infty}\frac{v_n}{n^{2n}}=\sqrt{\mathrm{e}}\,.$$

Finalement : $v_n\underset{n\to\infty}{\sim}\sqrt{\mathrm{e}}\,n^{2n}$.

Exemple

Soit f définie sur un intervalle I contenant 0 et continue en 0.
On suppose qu'il existe $k>1$ tel que le rapport $\frac{f(kx)-f(x)}{x}$ admette une limite finie ℓ en 0.
Montrer que f est dérivable en 0 et exprimer $f'(0)$ en fonction de k et ℓ.

L'hypothèse de l'énoncé peut s'écrire : $\frac{f(kx)-f(x)}{x}\underset{x\to0}{=}\ell+\varphi(x)$, *avec* $\lim\limits_{x\to0}\varphi(x)=0$, *ou encore :*

$$f(kx)-f(x)\underset{x\to0}{=}\ell x+x\varphi(x)\,.$$

Pour $x\in I$ *et pour tout* $n\in\mathbb{N}^*$, $\frac{x}{k^{n-1}}$ *appartient à* I *(car* $k>1$*) et l'on a :*

$$f\left(\frac{x}{k^{n-1}}\right)-f\left(\frac{x}{k^n}\right)=\ell\frac{x}{k^n}+\frac{x}{k^n}\varphi\left(\frac{x}{k^n}\right).$$

En sommant ces relations pour n *variant de* 1 *à* $N\in\mathbb{N}^*$, *on obtient, après télescopage :*

$$f(x)-f\left(\frac{x}{k^N}\right)=\ell x\sum_{n=1}^{N}\frac{1}{k^n}+\sum_{n=1}^{N}\frac{x}{k^n}\varphi\left(\frac{x}{k^n}\right)\quad(*)$$

φ *ayant une limite en* 0 *est bornée sur un intervalle* J *contenant* 0. *Pour* $x\in J$, *la série de terme général* $\frac{x}{k^n}\varphi\left(\frac{x}{k^n}\right)$ *est alors absolument convergente puisque* $\left|\frac{x}{k^n}\varphi\left(\frac{x}{k^n}\right)\right|\leqslant\frac{|x|\,\|\varphi\|_\infty}{k^n}$, *et puisque la série* $\sum\frac{1}{k^n}$ *est une série géométrique convergente (car* $k>1$*).*
De plus, $\lim\limits_{N\to+\infty}f\left(\frac{x}{k^N}\right)=f(0)$ *puisque* $k>1$ *et* f *continue en* 0, *et on sait que* $\sum\limits_{n=1}^{+\infty}\frac{1}{k^n}=\frac{1}{k-1}$. *On obtient donc, en faisant* $N\to+\infty$ *dans* $(*)$ *:*

$$\forall\,x\in J,\ f(x)-f(0)=\frac{\ell x}{k-1}+\sum_{n=1}^{+\infty}\frac{x}{k^n}\varphi\left(\frac{x}{k^n}\right),$$

soit encore :

$$\forall\, x \in J \setminus \{0\}\,,\ \frac{f(x)-f(0)}{x} = \frac{\ell}{k-1} + \sum_{n=1}^{+\infty} \frac{1}{k^n}\varphi\left(\frac{x}{k^n}\right).$$

Remarquons enfin que la série de fonctions $\left(x \mapsto \frac{1}{k^n}\varphi\left(\frac{x}{k^n}\right)\right)$ *est normalement convergente sur* J *(on majore* $\left|\varphi\left(\frac{x}{k^n}\right)\right|$ *uniformément par* $\|\varphi\|_\infty$*), elle y est donc uniformément convergente et le théorème de la double limite s'applique :*

$$\lim_{x\to 0}\left(\sum_{n=1}^{+\infty} \frac{1}{k^n}\varphi\left(\frac{x}{k^n}\right)\right) = \sum_{n=1}^{+\infty} \lim_{x\to 0}\left(\frac{1}{k^n}\varphi\left(\frac{x}{k^n}\right)\right) = 0\,,$$

puisque $\lim\limits_{x\to 0}\varphi(x) = 0$.

En conclusion, $\lim\limits_{x\to 0}\dfrac{f(x)-f(0)}{x} = \dfrac{\ell}{k-1}$*, ce qui prouve que* f *est dérivable en* 0 *et que* $f'(0) = \dfrac{\ell}{k-1}$.

✓ Notez dans les deux exemples précédents l'intérêt d'écrire les définitions de la dérivabilité et de la limite à l'aide d'un développement limité. De plus, il est plus commode d'écrire ces développements limités à l'aide d'une fonction φ qui tend vers 0 plutôt qu'avec des « o » ; cela permet de les manipuler bien plus facilement.

SF9.2 Utiliser le théorème de prolongement de la dérivée

Ce théorème est malheureusement souvent oublié. Il peut bien sûr servir lors de l'étude d'une fonction, pour examiner la dérivabilité en un point litigieux, mais aussi dans des exercices plus théoriques.

Exemple

Montrer que la fonction $f : \left\{\begin{array}{ccc} \left]0\,;\frac{\pi}{2}\right] & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{1}{x} - \dfrac{1}{\sin x} \end{array}\right.$ se prolonge en une fonction de classe $\mathscr{C}^1$ sur $\left[0\,;\frac{\pi}{2}\right]$.

Notons déjà que, en vertu des théorèmes usuels, f *est de classe* $\mathscr{C}^\infty$ *sur* $\left]0\,;\frac{\pi}{2}\right]$.

- *Pour* $x \neq 0$*, on a* $f(x) = \dfrac{\sin x - x}{x\sin x}$*. En utilisant le développement limité de* sin *au voisinage de* 0*, on a immédiatement :*

$$f(x) \underset{x\to 0^+}{\sim} \frac{-\frac{x^3}{6}}{x^2} = -\frac{x}{6}\,.$$

On en déduit $\lim\limits_{x\to 0^+} f(x) = 0$*, donc* f *se prolonge par continuité en* 0*, en posant* $f(0) = 0$.
On notera encore f *la fonction ainsi prolongée sur* $\left[0\,;\frac{\pi}{2}\right]$.

✓ *L'équivalent ci-dessus donne aussi le développement limité d'ordre* 1 *:* $f(x) = -\dfrac{x}{6} + o(x)$. *Cela montre tout de suite que* f *est dérivable en* 0 *et que* $f'(0) = -\dfrac{1}{6}$. *Mais cela ne suffit pas pour répondre à la question.*

- *Pour* $x \neq 0$ *on a :*

$$f'(x) = -\frac{1}{x^2} + \frac{\cos x}{\sin^2 x} = \frac{x^2\cos x - \sin^2 x}{x^2 \sin^2 x} \underset{x\to 0}{\sim} \frac{x^2\cos x - \sin^2 x}{x^4}\,.$$

On effectue un développement limité rapide :

$$\begin{aligned} x^2\cos x - \sin^2 x &\underset{x\to 0}{=} x^2\left(1 - \frac{x^2}{2} + o\big(x^2\big)\right) - \left(x - \frac{x^3}{6} + o\big(x^3\big)\right)^2 \\ &\underset{x\to 0}{=} x^2 - \frac{x^4}{2} - x^2 + 2\frac{x^4}{6} + o\big(x^4\big) = -\frac{x^4}{6} + o\big(x^4\big)\,, \end{aligned}$$

donc $\lim\limits_{x\to 0^+} f'(x) = -\dfrac{1}{6}$.

Le théorème de prolongement de la dérivée permet alors d'affirmer que f *est dérivable en* 0^+, *et que* $f'_d(0) = \lim\limits_{x\to 0^+} f'(x) = -\dfrac{1}{6}$. f' *est donc continue en* 0, *et par suite* f *est de classe* $\mathscr{C}^1$ *sur* $\left[0\,;\frac{\pi}{2}\right]$.

✓ Savoir utiliser à bon escient un développement limité, et savoir le calculer, fait partie des techniques de base de 1re année qui doivent être acquises.

Exemple

Soit f continue sur $[0\,;1]$ et dérivable sur $]0\,;1[$, telle que $\lim\limits_{x\to 0^+} xf'(x) = \ell$ existe.

Montrer que $\ell = 0$.

Soit $g\colon x \mapsto xf(x)$. g *est elle aussi continue sur* $[0\,;1]$ *et dérivable sur* $]0\,;1[$, *et pour tout* $x \in]0\,;1[$ *on a* $g'(x) = f(x) + xf'(x)$.
Les hypothèses de l'énoncé impliquent : $\lim\limits_{x\to 0^+} g'(x) = f(0) + \ell$.
Le théorème de prolongement de la dérivée implique alors que g *est dérivable à droite en* 0 *et que* $g'_d(0) = f(0) + \ell$.
Mais la définition de la dérivée donne :

$$g'_d(0) = \lim_{x\to 0^+} \frac{g(x) - g(0)}{x - 0} = \lim_{x\to 0^+} \frac{g(x)}{x} = \lim_{x\to 0^+} f(x) = f(0)\,,$$

donc $f(0) = f(0) + \ell$ *d'où* $\ell = 0$.

SF9.3 Utiliser le théorème de Rolle

On pense bien sûr à utiliser le théorème de Rolle dans tout problème d'analyse où interviennent les zéros d'une fonction.

Exemple

Soit $n \in \mathbb{N}^*$. Montrer que la fonction : $f\colon x \mapsto \left[(x^2-1)^n\right]^{(n)}$ possède exactement n zéros distincts dans l'intervalle $]-1\,;1[$.

Notons $g\colon x \mapsto (x^2-1)^n$, *de sorte que* $f = g^{(n)}$.

- *On remarque déjà que f est une fonction polynomiale de degré n, car dérivée n^e d'une fonction polynomiale de degré $2n$.*
- *-1 et 1 sont des racines d'ordre n du polynôme $(X^2-1)^n$. D'après les résultats du cours sur les polynômes, ce sont donc aussi des racines d'ordre $n-k$ $(0 \leqslant k \leqslant n-1)$ de sa dérivée k^e. En particulier, ce sont des zéros de $g, g', \ldots, g^{(n-1)}$.*
  - *Le théorème de Rolle appliqué à g entre -1 et 1 assure l'existence de x_{11} tel que $-1 < x_{11} < 1$ et $g'(x_{11}) = 0$.*
  - *Puis appliqué à g' entre -1 et x_{11} et entre x_{11} et 1, il donne l'existence de x_{21} et de x_{22} tels que $-1 < x_{21} < x_{11} < x_{22} < 1$ et $g''(x_{21}) = g''(x_{22}) = 0$.*

  En itérant le procédé, on obtient finalement que $g^{(n)}$ possède n racines distinctes dans $]-1\,;1[$. Puisqu'il s'agit d'une fonction polynomiale de degré n, il n'y en a pas d'autre.

Beaucoup d'autres résultats découlent de façon moins immédiate du théorème de Rolle. C'est le cas de la formule de Taylor-Lagrange, présentée dans l'exemple ci-dessous et qui généralise le théorème des accroissements finis.

Exemple

Soit f définie sur un intervalle I d'intérieur non vide, à valeurs réelles, de classe $\mathscr{C}^n$ et $n+1$ fois dérivable sur I.
Soient a et b deux points distincts de I. Montrer qu'il existe c strictement compris entre a et b tel que :

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c)\,.$$

Pour $x \in I$, posons :

$$\varphi(x) = f(b) - \sum_{k=0}^{n} \frac{(b-x)^k}{k!} f^{(k)}(x) - \lambda \frac{(b-x)^{n+1}}{(n+1)!}\,,$$

où λ est le réel tel que $\varphi(a) = 0$; un tel réel existe et est unique puisqu'il est obtenu en résolvant une équation du 1er degré de la forme $\lambda \dfrac{(b-a)^{n+1}}{(n+1)!} = cste$ avec $b - a \neq 0$.
Puisque f est de classe $\mathscr{C}^n$ et $n+1$ fois dérivable, φ est continue et dérivable sur I. De plus, $\varphi(a) = \varphi(b) = 0$, donc d'après le théorème de Rolle il existe c strictement compris entre a et b tel que $\varphi'(c) = 0$.
On calcule alors, pour tout $x \in I$:

$$\begin{aligned}
\varphi'(x) &= \sum_{k=1}^{n} \frac{(b-x)^{k-1}}{(k-1)!} f^{(k)}(x) - \sum_{k=0}^{n} \frac{(b-x)^k}{k!} f^{(k+1)}(x) + \lambda \frac{(b-x)^n}{n!} \\
&= \sum_{k=0}^{n-1} \frac{(b-x)^k}{k!} f^{(k+1)}(x) - \sum_{k=0}^{n} \frac{(b-x)^k}{k!} f^{(k+1)}(x) + \lambda \frac{(b-x)^n}{n!} \\
&= \left(\lambda - f^{(n+1)}(x)\right) \frac{(b-x)^n}{n!}\,.
\end{aligned}$$

La relation $\varphi'(c) = 0$ implique alors, puisque $b \neq c$, $\lambda = f^{(n+1)}(c)$. En réécrivant la relation $\varphi(a) = 0$ avec cette valeur de λ, on obtient la relation de l'énoncé.

✓ La méthode utilisée ici est classique : il s'agit d'appliquer le théorème de Rolle à une fonction auxiliaire « bien choisie ». La difficulté est bien sûr de trouver cette fonction...

SF9.4 Utiliser le théorème des accroissements finis

Il faut penser à utiliser le théorème des accroissements finis lorsque l'on rencontre une expression de la forme $f(x) - f(y)$ avec f dérivable.

Exemple

Calculer : $\lim\limits_{x\to+\infty} x^2\left(\mathrm{e}^{\frac{1}{x}} - \mathrm{e}^{\frac{1}{x+1}}\right)$.

La fonction $x \mapsto \mathrm{e}^{\frac{1}{x}}$ est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}_+^*$ et $f'(x) = -\dfrac{1}{x^2}\mathrm{e}^{\frac{1}{x}}$ pour tout $x > 0$.
En appliquant le théorème des accroissements finis à cette fonction sur $[x\,;x+1]$ avec $x > 0$ on obtient :

$$\exists\, c_x \in]x\,;x+1[,\ x^2\left(\mathrm{e}^{\frac{1}{x}} - \mathrm{e}^{\frac{1}{x+1}}\right) = \frac{x^2}{c_x^2}\mathrm{e}^{\frac{1}{c_x}}\,.$$

Puisque c_x est compris entre x et $x+1$, on a par le théorème d'encadrement $\lim\limits_{x\to+\infty} \dfrac{x^2}{c_x^2} = 1$ et $\lim\limits_{x\to+\infty} \mathrm{e}^{\frac{1}{c_x}} = 1$. La limite cherchée est donc égale à 1.

✓ Il ne faut pas oublier ; lors de l'application du théorème des accroissements finis, que le terme c dépend des bornes de l'intervalle. C'est pourquoi nous l'avons noté c_x dans l'exemple précédent, pour ne pas oublier qu'il dépend de x.

L'inégalité des accroissement finis est particulièrement utile pour reconnaître une fonction lipschitzienne. Nous nous en sommes déjà servi en **[SF8.7]** lors de l'utilisation du théorème du point fixe pour l'étude de suites récurrentes.

Exemple

Soit f dérivable de $\mathbb{R}_+$ dans $\mathbb{R}$.

1) Montrer que si f' est bornée sur $\mathbb{R}_+$, f est uniformément continue sur $\mathbb{R}_+$.

2) Montrer que si $\lim\limits_{x\to+\infty} |f'(x)| = +\infty$, f n'est pas uniformément continue sur $\mathbb{R}_+$.

1) *C'est un résultat connu. L'inégalité des accroissements finis implique :*

$$\forall\,(x,y) \in (\mathbb{R}_+)^2,\ |f(y) - f(x)| \leqslant \|f'\|_\infty\, |y - x|\,.$$

Ainsi, f est lipschitzienne sur $\mathbb{R}_+$ et donc y est uniformément continue.

2) *Supposons $\lim\limits_{x\to+\infty} |f'(x)| = +\infty$ et, par l'absurde, f uniformément continue sur $\mathbb{R}_+$.*
En prenant $\varepsilon = 1$ dans la définition, on obtient :

$$\exists\,\alpha > 0,\ \forall\,(x,y) \in (\mathbb{R}_+)^2,\ |x - y| \leqslant \alpha \Longrightarrow |f(x) - f(y)| \leqslant 1\,.$$

En particulier, pour tout $x \in \mathbb{R}$ on a $|f(x+\alpha) - f(x)| \leqslant 1$.
Or d'après le théorème des accroissements finis, il existe $c_x \in]x\,;x+\alpha[$ tel que $f(x+\alpha) - f(x) = \alpha f'(c_x)$.
On a donc $|f'(c_x)| \leqslant \dfrac{1}{\alpha}$; mais quand $x \to +\infty$, $c_x \to +\infty$, il y a donc contradiction avec l'hypothèse.
On a donc montré par l'absurde que f n'est pas uniformément continue sur $\mathbb{R}_+$.

SF9.5 Utiliser les variations

L'étude des variations d'une fonction auxiliaire pour résoudre un problème, en particulier pour démontrer une inégalité, est une méthode très courante et importante. La difficulté est parfois de trouver la « bonne » fonction auxiliaire.

Exemple

Soit $f\colon [a\,;b] \to \mathbb{R}$, dérivable, telle que $f(a) = 0$. On suppose qu'il existe $k \geqslant 0$ tel que :

$$\forall\, x \in [a\,;b],\ |f'(x)| \leqslant k|f(x)|\,.$$

Montrer que $f = 0$.

Considérons la fonction $g\colon x \mapsto \big(f(x)\big)^2 \mathrm{e}^{-2kx}$. f étant dérivable sur $[a\,;b]$, g l'est aussi et l'on a :

$$\forall\, x \in [a\,;b],\ g'(x) = 2\left(f'(x) - kf(x)\right) f(x)\mathrm{e}^{-2kx}.$$

Considérons deux cas.

- *Si $f(x) \geqslant 0$, l'hypothèse de l'énoncé implique $f'(x) - kf(x) \leqslant 0$, donc $g'(x) \leqslant 0$.*
- *Si $f(x) \leqslant 0$, l'inégalité $f'(x) \geqslant -k\,|f(x)|$ implique $f'(x) \geqslant kf(x)$, et là encore $g'(x) \leqslant 0$.*

On a donc $g' \leqslant 0$ sur $[a\,;b]$. g' est donc décroissante et, puisque $g(a) = 0$, $g(x) \leqslant 0$ pour tout x, ce qui implique $\big(f(x)\big)^2 \leqslant 0$ donc f est nulle.

SF9.6 Résoudre une équation fonctionnelle

Résoudre une équation fonctionnelle, c'est-à-dire trouver toutes les fonctions qui satisfont à certaines conditions, n'est jamais très facile. Une des solutions possibles est, après avoir montré la dérivabilité des fonctions concernées, d'essayer d'obtenir des conditions nécessaires sur la dérivée ou les dérivées successives.

Exemple

Trouver toutes les applications $f\colon \mathbb{R} \to \mathbb{R}$, dérivables, telles que :

$$\forall\,(x,y) \in \mathbb{R}^2,\ f(x) - f(y) = (x-y)f'\left(\frac{x+y}{2}\right).$$

Soit f une solution du problème.

- *En posant $z = \dfrac{x+y}{2}$, on a, pour tous y, z réels tels que $z \neq y$:*

$$f'(z) = \frac{f(2z-y) - f(y)}{2(z-y)}\,,$$

ce qui montre que f' est dérivable sur $\mathbb{R}\setminus\{y\}$. y étant quelconque, on obtient que f' est dérivable sur $\mathbb{R}$, donc f est deux fois dérivable.
En dérivant alors la relation donnée par rapport à x on obtient :

$$f'(x) = \frac{x-y}{2} f''\left(\frac{x+y}{2}\right) + f'\left(\frac{x+y}{2}\right),$$

ou encore, en posant $z = \dfrac{x+y}{2}$,

$$\forall\,(x,z)\in\mathbb{R}^2,\ f'(x) = (x-z)f''(z) + f'(z)\,.$$

En dérivant cette relation par rapport à x, on obtient $f''(x) = f''(z)$ pour tout $(x,z)\in\mathbb{R}^2$, c'est-à-dire que f'' est constante.

f est donc nécessairement de la forme : $f\colon x\mapsto ax^2+bx+c$.

- *Réciproquement, en remplaçant dans l'équation initiale, il est facile de vérifier que toutes ces fonctions sont bien solutions.*

■ Comment montrer et utiliser la convexité d'une fonction ?

Il y a plusieurs caractérisations des fonctions convexes. Les deux premières que nous présentons ci-après sont à utiliser si l'énoncé ne donne pas de renseignement quant à la dérivabilité.

SF9.7 Revenir à la définition

Une façon de retenir la définition d'une fonction convexe peut être la suivante :

« l'image du barycentre est inférieure au barycentre des images »

ou encore :

« la courbe est située sous ses cordes ».

Bien sûr, il faut ensuite savoir traduire ces phrases en propriétés mathématiques rigoureuses !

Exemple

Soit $f\colon I\to\mathbb{R}$ une fonction strictement convexe, et $x<y<z$ dans I.

Montrer que le déterminant $\begin{vmatrix} 1 & x & f(x)\\ 1 & y & f(y)\\ 1 & z & f(z)\end{vmatrix}$ est strictement positif.

Puisque $y\in\,]x\,;z[$, il existe $\lambda\in\,]0\,;1[$ tel que $y=\lambda x+(1-\lambda)z$. On a alors :

$$z-y=\lambda(z-x)\ ;\ y-x=(1-\lambda)(z-x)\ ;\ f(y)<\lambda f(x)+(1-\lambda)f(z)\,,$$

d'où, en développant le déterminant selon la 3e colonne :

$$\begin{aligned}\begin{vmatrix} 1 & x & f(x)\\ 1 & y & f(y)\\ 1 & z & f(z)\end{vmatrix} &= (z-y)f(x) - f(y)(z-x) + (y-x)f(z)\\ &= \lambda(z-x)f(x) - f(y)(z-x) + (1-\lambda)(z-x)f(z)\\ &= (z-x)\big(\lambda f(x) + (1-\lambda)f(z) - f(y)\big) > 0\,.\end{aligned}$$

Exemple

Soit $f : \mathbb{R} \to \mathbb{R}$.
Montrer que f est convexe si et seulement si pour tout segment $[a\,;b]$ et tout réel c, la restriction de $f - c\mathrm{Id}_{\mathbb{R}}$ à $[a\,;b]$ atteint son maximum en a ou en b.

- *Supposons f convexe sur $[a\,;b]$. Par définition, le graphe de la fonction est situé en dessous de ses cordes, ce qui signifie que, si l'on note $x \mapsto \ell(x)$ l'équation de la droite passant par les points de coordonnées $\big(a,f(a)\big)$ et $\big(b,f(b)\big)$, on a $f(x) \leqslant \ell(x)$ pour tout $x \in [a\,;b]$.*

  ✓ *En fait on a $\ell(x) = \dfrac{f(b)-f(a)}{b-a}(x-a) + f(a)$; cette expression nous servira dans la question suivante.*

  ℓ étant une fonction affine, son maximum est atteint en a ou en b, et il est égal à $\max\big(f(a),f(b)\big)$. Il en résulte que :

$$\max_{x\in[a;b]} f(x) = \max\big(f(a),f(b)\big)\,.$$

  Pour obtenir le résultat demandé, il suffit alors de remarquer que, si f est convexe, il en est de même de la fonction $f - c\mathrm{Id}_{\mathbb{R}}$ pour tout réel c (car somme de deux fonctions convexes).

- *Réciproquement, supposons la propriété de l'énoncé vérifiée, et montrons que f est convexe. Pour cela nous allons raisonner par contraposition.*

  Si f n'est pas convexe, il existe un point du segment ouvert $]a\,;b[$ où le point de la courbe est strictement au-dessus du point correspondant de la corde, c'est-à-dire que, avec les notations précédentes :

$$\exists\, t \in]a\,;b[,\ f(t) > \ell(t)\,.$$

  En choisissant $c = \dfrac{f(b)-f(a)}{b-a}$, on a donc :

$$f(t) - ct > \ell(t) - ct = f(a) - ca = f(b) - cb\,,$$

  ce qui signifie que $f - c\mathrm{Id}_{\mathbb{R}}$ n'atteint son maximum ni en a ni en b.

  Il existe donc un c réel pour lequel la propriété de l'énoncé n'est pas vérifiée, ce qui démontre le résultat cherché par contraposition.

SF9.8 Utiliser la croissance de la pente

Exemple

Soit $f : \mathbb{R} \to \mathbb{R}$, convexe et majorée. Montrer que f est constante.

On sait que pour tout $a \in \mathbb{R}$, l'application $F_a : x \mapsto \dfrac{f(x)-f(a)}{x-a}$ est croissante sur $\mathbb{R}\setminus\{a\}$.
Pour tout $x \geqslant a+1$ on a, en notant M un majorant de f :

$$F_a(x) \leqslant \frac{M-f(a)}{x-a} \leqslant M - f(a)\,,$$

donc F_a est majorée sur $[a+1\,;+\infty[$. D'après le théorème de la limite monotone, elle possède une limite en $+\infty$, et

$$\lim_{x\to+\infty} F_a(x) \leqslant \lim_{x\to+\infty} \frac{M-f(a)}{x-a} = 0\,.$$

On en déduit que $F_a(x) \leqslant 0$ pour tout $x > a$, et cela étant vrai pour tout $a \in \mathbb{R}$, on en déduit que f est décroissante sur $\mathbb{R}$.

Étant décroissante majorée, elle admet une limite en $-\infty$, ce qui implique $\lim\limits_{x\to-\infty} F_a(x) = 0$. F_a étant croissante, elle est positive, ce qui implique f croissante.

En conclusion, f est décroissante et croissante, elle est constante.

SF9.9 Utiliser les dérivées

Cette méthode ne peut évidemment s'utiliser que pour une fonction dérivable ou deux fois dérivable. Elle est très commode pour les fonctions usuelles.

Exemple

Montrer que la fonction $f : x \mapsto \ln(1+\mathrm{e}^x)$ est strictement convexe sur $\mathbb{R}$.

f étant de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$, il n'y a qu'à calculer, pour tout x :

$$f'(x) = \frac{\mathrm{e}^x}{1+\mathrm{e}^x} = 1 - \frac{1}{1+\mathrm{e}^x}\,,\ f''(x) = \frac{\mathrm{e}^x}{(1+\mathrm{e}^x)^2}\,\cdot$$

Donc f'' est strictement positive sur $\mathbb{R}$ et f est strictement convexe.

On peut aussi utiliser les propriétés de la dérivée dans des exercices moins calculatoires.

Exemple

Soit $f : \mathbb{R}_+ \to \mathbb{R}$, dérivable et concave, telle que $f(0) \geqslant 0$.

Montrer que : $\forall\,(x,y) \in (\mathbb{R}_+)^2,\ f(x+y) \leqslant f(x) + f(y)\,.$

Pour tout $y \geqslant 0$, considérons la fonction $\varphi : x \in \mathbb{R}_+ \mapsto f(x+y) - f(x) - f(y)$. Puisque f est dérivable, φ l'est et ;

$$\forall\, x \in \mathbb{R}_+,\ \varphi'(x) = f'(x+y) - f'(x)\,.$$

Puisque f est concave, sa dérivée est décroissante, donc $\varphi' \leqslant 0$.
Ainsi φ est décroissante, et puisque $\varphi(0) = -f(0) \leqslant 0$, φ est négative sur $\mathbb{R}_+$, ce qui donne l'inégalité demandée.

✓ Remarquez que nous avons utilisé ici une méthode simple et très classique pour démontrer une inégalité : l'étude de la fonction différence.

SF9.10 Utiliser la position par rapport à la tangente

Cette méthode est surtout utilisée pour démontrer certaines inégalités.

Exemple

Montrer que : $\forall\, x \in \left[0\,;\tfrac{\pi}{2}\right],\ \dfrac{2}{\pi}x \leqslant \sin x \leqslant x\,.$

La fonction sin *est concave sur $\left[0\,;\tfrac{\pi}{2}\right]$ (dérivée seconde négative) donc sa courbe est située en-dessous de sa tangente à l'origine qui est la droite d'équation $y = x$, et au-dessus de la corde qui joint les points de coordonnées $(0,0)$ et $\left(\tfrac{\pi}{2},1\right)$, qui est la droite d'équation $y = \dfrac{2}{\pi}x$.*

✓ On montre de la même façon deux autres inégalités classiques :

$$\forall x > -1,\ \ln(1+x) \leqslant x \quad \text{et} \quad \forall x \in \mathbb{R},\ \mathrm{e}^x \geqslant 1+x\,.$$

SF9.11 Démontrer ou utiliser une inégalité de convexité

Nous venons de voir certaines inégalités de convexité qui s'obtiennent grâce à la position de la courbe par rapport à sa tangente. D'autres s'obtiennent grâce aux barycentres.

Exemple

En utilisant la convexité de la fonction $x \mapsto \ln(1+\mathrm{e}^x)$, démontrer :

$$\forall n \in \mathbb{N}^*,\ \forall (x_1,\dots,x_n) \in (\mathbb{R}_+^*)^n,\quad 1+\left(\prod_{k=1}^{n} x_k\right)^{1/n} \leqslant \left(\prod_{k-1}^{n}(1+x_k)\right)^{1/n}.$$

La convexité de la fonction $f\colon x \mapsto \ln(1+\mathrm{e}^x)$ a été démontrée dans un exemple précédent, par le simple calcul de f''.
On a donc, pour tout $(a_1,\dots,a_n) \in \mathbb{R}^n$:

$$f\left(\frac{a_1+\cdots+a_n}{n}\right) \leqslant \frac{f(a_1)+\cdots+f(a_n)}{n}\,,$$

soit :

$$\ln\left(1+\mathrm{e}^{\frac{a_1+\cdots+a_n}{n}}\right) \leqslant \frac{1}{n}\ln\left(\prod_{k=1}^{n}(1+\mathrm{e}^{a_k})\right),$$

puis :

$$1+\mathrm{e}^{\frac{a_1+\cdots+a_n}{n}} \leqslant \left(\prod_{k=1}^{n}(1+\mathrm{e}^{a_k})\right)^{1/n}.$$

Il suffit d'appliquer cette inégalité à $a_k = \ln x_k$ pour obtenir l'inégalité de l'énoncé.

À vous de jouer !

Exercice 9.1 **SF9.3**

Soit $f[a\,;b] \to \mathbb{R}$, de classe $\mathscr{C}^n$ $(n \in \mathbb{N}^*)$.
On suppose que f s'annule en n points distincts $a_1 < a_2 < \cdots < a_n$. Montrer que pour tout $x \in [a\,;b]$, il existe $c \in [a\,;b]$ tel que :

$$f(x) = \frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{n!} f^{(n)}(c)\,.$$

Exercice 9.2 **SF9.3**

Soient p et q deux fonctions continues sur un segment $[a\,;b]$ à valeurs réelles.

Montrer que si f est une solution de l'équation différentielle : $y''+p(x)y'+q(x)y=0$ qui s'annule une infinité de fois sur $[a\,;b]$, alors f est nulle sur $[a\,;b]$.

Exercice 9.3 **SF9.5** **SF9.9** **SF9.10**

Soit f une application définie sur $\mathbb{R}_+$, à valeurs positives, deux fois dérivable et majorée. On suppose qu'il existe un réel $a \geqslant 0$ tel que : $\forall\, x \in \mathbb{R}_+,\ f''(x) \geqslant af(x)$.

1) Montrer que f est décroissante, puis que : $\lim\limits_{x\to+\infty} f(x) = \lim\limits_{x\to+\infty} f'(x) = 0$.

2) Montrer que : $\forall\, x \geqslant 0,\ f(x) \leqslant f(0)\mathrm{e}^{-x\sqrt{a}}$.

Exercice 9.4 **SF9.7**

Soit $f\colon I \to \mathbb{R}$ une fonction convexe. Montrer que si $a < b < c$ sont dans I alors :

$$f(a-b+c) \leqslant f(a) - f(b) + f(c)\,.$$

Exercice 9.5 **SF9.4** **SF9.9**

Soit f dérivable et convexe de $\mathbb{R}_+$ dans $\mathbb{R}$.
Pour tout $x \geqslant 0$ on pose : $g(x) = xf'(x) - f(x)$. Montrer que g est croissante.

Exercice 9.6 **SF9.8**

Soit f une application convexe de $[a\,;b]$ dans $\mathbb{R}$.

1) On suppose qu'il existe $c \in\,]a\,;b[$ tel que $f(a) = f(c) = f(b)$.
Montrer que f est constante.

2) On suppose qu'il existe $c \in\,]a\,;b[$ tel que $\dfrac{f(b)-f(a)}{b-a} = \dfrac{f(c)-f(a)}{c-a}$.
Montrer que f est affine.

Thème 10 - Dérivation et intégration des fonctions vectorielles

En ce qui concerne l'intégration, ce chapitre est essentiellement un chapitre de révisions sur le programme de 1re année. Cependant, certains exemples ou exercices proposés peuvent aussi faire appel à des notions qui seront détaillées dans le chapitre consacré aux intégrales généralisées.

■ Comment étudier un arc paramétré ?

SF10.1 Suivre le plan d'étude général

Soit un arc paramétré, défini par $\overrightarrow{OM}(t) = x(t)\,\vec{i} + y(t)\,\vec{j}$, pour t appartenant à une partie D de $\mathbb{R}$.
On en donne ici un plan d'étude général.

1. **On cherche un domaine d'étude réduit.**
   Sans vouloir être exhaustif, citons ici les cas les plus courants ; ces cas sont aisés à retrouver à l'aide d'une petite figure et/ou un peu de réflexion.
   - Si x *et* y sont simultanément périodiques de période $T > 0$, on obtiendra toute la courbe en se limitant à un intervalle de longueur T.
   - Si x *et* y sont paires, on obtiendra toute la courbe en restreignant l'intervalle d'étude à $D \cap \mathbb{R}_+$.
   - Si x est paire *et* y impaire, on restreindra l'étude à $D \cap \mathbb{R}_+$ puis on complètera la courbe obtenue par une symétrie d'axe Ox.
   - Si x est impaire *et* y paire, on restreindra l'étude à $D \cap \mathbb{R}_+$ puis on complètera la courbe obtenue par une symétrie d'axe Oy.
   - Si x *et* y sont impaires, on restreindra l'étude à $D \cap \mathbb{R}_+$ puis on complètera la courbe obtenue par une symétrie par rapport au point O.
2. **On dresse le tableau de variations** de x et y sur ce domaine d'étude.
3. **On étudie les branches infinies éventuelles.**
   Il y a une branche infinie quand x ou/et y tend vers ∞ lorsque t tend vers une certaine valeur t_0 (éventuellement $\pm\infty$). Plusieurs cas sont possibles.
   - Si $\lim\limits_{t\to t_0} x(t) = \ell$ et $\lim\limits_{t\to t_0} y(t) = \pm\infty$, il y a une asymptote d'équation $x = \ell$.
   - Si $\lim\limits_{t\to t_0} y(t) = \ell$ et $\lim\limits_{t\to t_0} x(t) = \pm\infty$, il y a une asymptote d'équation $y = \ell$.
   - Si x et y tendent tous deux vers l'infini, une méthode possible d'étude est la suivante.
     On forme le rapport $\dfrac{y(t)}{x(t)}$ et on étudie sa limite quand $t \to t_0$.
     - Si cette limite n'existe pas, on ne peut rien dire.

- Si $\lim\limits_{t\to t_0} \dfrac{y(t)}{x(t)} = \infty$, il y a une branche parabolique de direction Oy.
- Si $\lim\limits_{t\to t_0} \dfrac{y(t)}{x(t)} = 0$, il y a une branche parabolique de direction Ox.
- Si $\lim\limits_{t\to t_0} \dfrac{y(t)}{x(t)} = a$, il y a une direction asymptotique de pente a et on considère alors la différence $y(t) - ax(t)$.
  - Si $y(t) - ax(t)$ n'a pas de limite quand $t \to t_0$, on ne peut rien dire.
  - Si $\lim\limits_{t\to t_0} y(t) - ax(t) = \infty$, il y a une branche parabolique de pente a.
  - Enfin, si $\lim\limits_{t\to t_0} y(t) - ax(t) = b$, il y a une asymptote d'équation $y = ax + b$.

4. **On étudie les points stationnaires éventuels.**
   Il faut se rappeler que la tangente en un point non stationnaire de paramètre t_0 a pour pente $\dfrac{y'(t_0)}{x'(t_0)}$. En un point stationnaire (c'est-à-dire tel que $x'(t_0) = y'(t_0) = 0$), pour déterminer la pente de la tangente, on peut chercher $\lim\limits_{t\to t_0} \dfrac{y'(t)}{x'(t)}$.
5. **On peut alors faire le tracé.**
   Vous pouvez vous contenter d'un tracé approximatif (mais pas trop !).
6. **Si cela vous est demandé** ou s'il vous reste du temps, on peut éventuellement déterminer les éventuels points multiples. Pour cela, il faut résoudre le système :
$$\begin{cases} x(t) = x(u) \\ y(t) = y(u) \end{cases} \quad \text{avec} \quad t \neq u\,.$$

Exemple

Étudier et tracer la courbe paramétrée par :
$$\begin{cases} x(t) = \dfrac{t}{1+t^3} \\ y(t) = \dfrac{t^2}{1+t^3} \end{cases} \quad \text{pour } t \in \mathbb{R}\,.$$

Les applications x et y dont définies et de classe $\mathscr{C}^\infty$ sur $\mathbb{R} \setminus \{-1\}$ et pour tout $t \neq -1$ on a :
$$x'(t) = \frac{1-2t^3}{(1+t^3)^2} \quad \text{et} \quad y'(t) = \frac{t(2-t^3)}{(1+t^3)^2}\,.$$
Il ne reste plus qu'à dresser le tableau de variations, ce qui ne pose pas de problèmes ; vous noterez la disposition de ce tableau : nous avons regroupé les variations de x et de y l'une en dessous de l'autre car cela permet à notre avis de mieux visualiser la courbe.

t	$-\infty$		-1		0		$\frac{1}{\sqrt[3]{2}}$		$\sqrt[3]{2}$		$+\infty$
$x'(t)$	$+$		‖	$+$	$+$	0	$-$		$-$		
$x(t)$	0	↗	$+\infty$ ‖ $-\infty$	↗	0	↗	$\approx 0{,}53$	↘	$\approx 0{,}42$	↘	0
$y(t)$	0	↘	$-\infty$ ‖ $+\infty$	↘	0	↗	$\approx 0{,}42$	↗	$\approx 0{,}53$	↘	0
$y'(t)$	$-$		‖	$-$	0	$+$	$+$	0	$-$		

On remarque qu'il y a une branche infinie lorsque t tend vers -1^+ et vers -1^-. Étudions la, en calculant :

$$\lim_{t\to -1} \frac{y(t)}{x(t)} = \lim_{t\to -1} t = -1 \text{ puis } \lim_{t\to -1} \big(y(t)+x(t)\big) = \lim_{t\to -1} \frac{t}{1-t+t^2} = -\frac{1}{3}\,.$$

Cela montre que la droite d'équation $y = -x - \frac{1}{3}$ est asymptote à la courbe.

De plus, $y(t) + x(t) + \frac{1}{3} = \frac{(1+t)^2}{3(1-t+t^2)}$ est toujours positif, donc la courbe est située au-dessus de son asymptote.

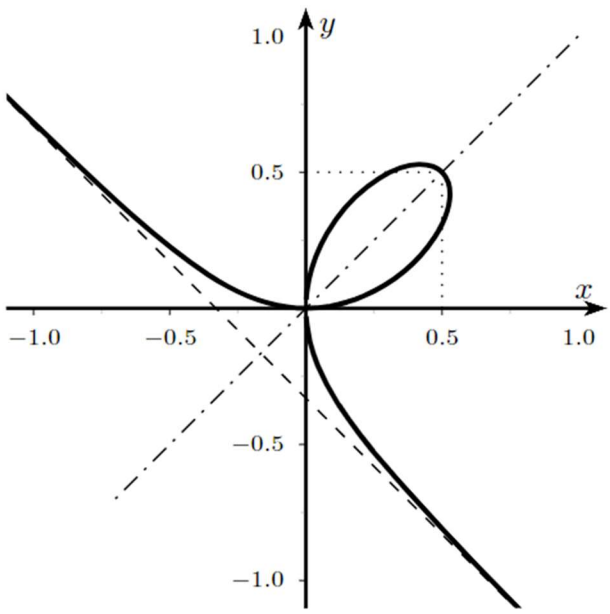
Cette courbe s'appelle un folium de Descartes, *elle est représentée ci-dessous.*

On note alors qu'il semble y avoir une symétrie par rapport à la droite d'équation $y = x$. Démontrons-le.

Pour cela, on remarque que pour tout $t \neq -1$ on a

$$x\left(\frac{1}{t}\right) = y(t) \quad \text{et} \quad y\left(\frac{1}{t}\right) = y(t)\,.$$

Ainsi, le point de paramètre $\frac{1}{t}$ se déduit de celui de paramètre t par la symétrie orthogonale par rapport à la 1re bissectrice. On aurait donc pu se contenter d'étudier les variations pour $t \in]-1\,;1]$.



■ Comment calculer une primitive classique ?

Il faut déjà bien connaître les primitives des fonctions usuelles, qui sont rappelées en **[S10.13]**.
Nous exposons ci-dessous des méthodes qui permettent de calculer une primitive dans les cas courants ; il s'agit des méthodes générales, qu'on peut appliquer de façon presque automatique, mais il ne faut pas oublier que dans de nombreux cas, un peu de réflexion peut beaucoup simplifier les calculs.
De plus nous nous contentons d'exposer ici des *techniques* de calcul. Aucun problème théorique n'est soulevé. C'est à vous qu'il conviendra, dans chaque cas, de préciser le domaine de validité des calculs (et aussi, dans le cas d'une intégrale généralisée, de prouver la convergence).

SF10.2 Calculer une primitive d'une fonction comportant polynômes et exponentielles

- *Primitives de la forme* $\int \mathrm{e}^{ax}P(x)\,\mathrm{d}x$ *où* $a \in \mathbb{C}$ *et* $P \in \mathbb{C}_n[X]$
  Il suffit de faire $n-1$ intégrations par parties successives, en dérivant à chaque fois le polynôme. On aura donc à la fin un résultat de la forme $\int \mathrm{e}^{ax}P(x)\,\mathrm{d}x = \mathrm{e}^{ax}Q(x)$, où Q est aussi un polynôme de degré n.

- *Primitives de la forme* $I = \int \mathrm{e}^{ax}\cos(bx)P(x)\,\mathrm{d}x$ *et* $J = \int \mathrm{e}^{ax}\sin(bx)P(x)\,\mathrm{d}x$ *où* $a, b \in \mathbb{R}$ *et* $P \in \mathbb{C}_n[X]$
  La méthode générale consiste à calculer :
  $$K = I + \mathrm{i}J = \int \mathrm{e}^{(a+\mathrm{i}b)x}P(x)\,\mathrm{d}x\,.$$
  ce qui ramène au cas précédent. On obtient alors I et J en prenant les parties réelle et imaginaire de la primitive trouvée.

  Dans le cas particulier $a = 0$, on peut calculer I ou J par $n-1$ intégrations par parties successives pour abaisser le degré du polynôme P.

- *Primitives de la forme* $I = \int P(x)\mathrm{e}^{ax}Q(\sin x, \cos x)\,\mathrm{d}x$ *où* $a \in \mathbb{R}$, $P \in \mathbb{C}_n[X]$, Q *polynôme à 2 indéterminées*
  La méthode générale consiste ici à linéariser l'expression $Q(\sin x, \cos x)$; on se ramène ainsi à des sommes de primitives du cas précédent.

  ✓ Nous vous renvoyons à votre cours de 1re année pour les méthodes de linéarisation d'un polynôme en sin et cos (ces techniques utilisent les nombres complexes en général).

SF10.3 Calculer une primitive d'une fonction rationnelle

Pour calculer une primitive $\int R(x)\,\mathrm{d}x$ où R est fonction rationnelle de $\mathbb{R}(X)$, le principe de base est de la décomposer en éléments simples de $\mathbb{R}(X)$.

- La partie entière (éventuelle) est un polynôme, qui s'intègre sans problème.
- Les éléments simples de première espèce, de la forme $\dfrac{1}{(x-a)^n}$ s'intègrent aussi directement.
- Pour les éléments simples de seconde espèce, de la forme $\dfrac{ax+b}{(x^2+px+q)^n}$, on commence par faire apparaître au numérateur la dérivée du dénominateur :

$$\frac{ax+b}{(x^2+px+q)^n} = \frac{a}{2}\frac{2x+p}{(x^2+px+q)^n} + \left(b - \frac{ap}{2}\right)\frac{1}{(x^2+px+q)^n}\,.$$

Le premier terme, qui est de la forme $\dfrac{u'}{u^n}$ s'intègre aisément.

Sachant que le trinôme x^2+px+q est irréductible, on peut l'écrire sous forme canonique : $x^2+px+q = (x-\alpha)^2+\beta^2$.

Dans l'intégrale $\displaystyle\int \frac{1}{[(x-\alpha)^2+\beta^2]^n}\,\mathrm{d}x$, le changement de variable $t = \dfrac{x-\alpha}{\beta}$ nous ramène au calcul d'intégrales de la forme :

$$I_n = \int \frac{\mathrm{d}t}{(t^2+1)^n}\,.$$

On connaît I_1, et on peut ensuite obtenir une relation de récurrence entre I_n et I_{n+1}, ce qui permettra de calculer I_n, connaissant $I_1 = \operatorname{Arctan} t$. Cette relation de récurrence s'obtient facilement en intégrant I_n par parties :

$$I_n = \int \frac{\mathrm{d}t}{(t^2+1)^n} = \frac{t}{(t^2+1)^n} + 2n\int \frac{t^2}{(t^2+1)^{n+1}}\,\mathrm{d}t\,,$$

et, pour l'intégrale de droite, on utilise l'astuce habituelle consistant à écrire $t^2 = (t^2-1)+1$, ce qui donne $I_n = \dfrac{t}{(t^2+1)^n} + 2n(I_n - I_{n+1})$.

SF10.4 Calculer une primitive d'une fonction avec des radicaux

On procède à un changement de variable « astucieux » pour se ramener à une fraction rationnelle classique que l'on décomposera ensuite en éléments simples.

- *Primitives de la forme* $\displaystyle\int R\left(x, \left(\frac{ax+b}{cx+d}\right)^{p/q}\right)\mathrm{d}x$ *où* R *est une fraction rationnelle à 2 indéterminées,* $(a,b,c,d) \in \mathbb{R}^4$ *et* $(p,q) \in \mathbb{N}\times\mathbb{N}^*$.

  On effectue le changement de variable $t = \left(\dfrac{ax+b}{cx+d}\right)^{1/q}$, qui conduit à

$$x = \frac{dt^q - b}{a - ct^q} \quad \text{et} \quad \mathrm{d}x = \frac{q(ad-bc)t^{q-1}}{(a-ct^q)^2}\,\mathrm{d}t\,,$$

  et on se ramène ainsi au calcul d'une primitive d'une fraction rationnelle en t.
- *Primitives de la forme* $\displaystyle\int R(x,\sqrt{ax+b})\,\mathrm{d}x$ *où* R *est une fraction rationnelle à deux indéterminées et* $(a,b) \in \mathbb{R}^2$.

  C'est un simple cas particulier (mais fréquent) du précédent ; il convient ici de poser $t = \sqrt{ax+b}$.
- *Primitives de la forme* $\displaystyle\int R\left(x,\sqrt{ax^2+bx+c}\right)\mathrm{d}x$ *où* R *est une fraction rationnelle à 2 indéterminées,* $(a,b,c) \in \mathbb{R}^3$ *tels que* $a \neq 0$ *et* $\Delta = b^2-4ac \neq 0$.

  On écrit le trinôme sous forme canonique :

$$(ax^2+bx+c = a\left(\left(x+\frac{b}{2a}\right)^2-\frac{\Delta}{4a^2}\right),$$

puis on effectue le changement de variable : $t = \dfrac{2ax+b}{\sqrt{|\Delta|}}$, ce qui permet de se ramener, selon les signes de a et Δ, à l'un des cas suivants :

$$\Delta > 0 \text{ et } a < 0 \quad : \quad I_1 = \int S\left(t, \sqrt{1-t^2}\right) \mathrm{d}t$$
$$\Delta > 0 \text{ et } a > 0 \quad : \quad I_2 = \int S\left(t, \sqrt{t^2-1}\right) \mathrm{d}t$$
$$\Delta < 0 \text{ et } a > 0 \quad : \quad I_3 = \int S\left(t, \sqrt{t^2+1}\right) \mathrm{d}t$$

où S est une fraction rationnelle à deux indéterminées.
On effectue alors un des changements de variable suivants :

$$\text{dans } I_1 \quad : \quad \text{poser } t = \sin u \text{ ou } t = \cos u\,,$$
$$\text{dans } I_2 \quad : \quad \text{poser } t = \pm \operatorname{ch} u \text{ ou } t = \frac{1}{\cos u}\,,$$
$$\text{dans } I_3 \quad : \quad \text{poser } t = \operatorname{sh} u \text{ ou } t = \tan u\,,$$

le choix dépendant souvent du domaine sur lequel on intègre.

SF10.5 Calculer une primitive d'une fonction rationnelle en sin et cos

Il s'agit ici de calculer une intégrale de la forme $\int R(\sin x, \cos x)\,\mathrm{d}x$ où R est une fraction rationnelle à deux indéterminées.

- La méthode générale (qui n'est pas toujours la plus économe en calcul) consiste à effectuer le changement de variable $t = \tan\frac{x}{2}$ (*attention au domaine !*). Avec les relations bien connues :
$$\cos(x) = \frac{1-t^2}{1+t^2} \quad ; \quad \sin(x) = \frac{2t}{1+t^2} \quad ; \quad \tan(x) = \frac{2t}{1-t^2}$$
et avec la relation $x = 2\operatorname{Arctan} t$ (lorsque $x \in]-\pi\,;\pi[$) qui conduit à $\mathrm{d}x = \dfrac{2\,\mathrm{d}t}{1+t^2}$, le calcul de l'intégrale $\int R(\sin x, \cos x)\,\mathrm{d}x$ se ramène à celui d'une fraction rationnelle en t.
- D'autres changements de variable sont possibles dans des cas particuliers.
Ce sont les *règles de Bioche.*
  - Si la *différentielle* $R(\sin x, \cos x)\,\mathrm{d}x$ est invariante par le changement $x \longmapsto -x$, on peut poser $t = \cos x$.
  - Si la *différentielle* $R(\sin x, \cos x)\,\mathrm{d}x$ est invariante par le changement $x \longmapsto \pi - x$, on peut poser $t = \sin x$.
  - Si la *différentielle* $R(\sin x, \cos x)\,\mathrm{d}x$ est invariante par le changement $x \longmapsto \pi + x$, on peut poser $t = \tan x$.
  - Si la *différentielle* $R(\sin x, \cos x)\,\mathrm{d}x$ est invariante par deux des changements de variable précédents, on peut poser $t = \cos 2x$.

SF10.6 Calculer une primitive d'une fonction rationnelle en sh, ch, exp

Pour calculer une intégrale de la forme $\int R(\operatorname{sh} x, \operatorname{ch} x)\,\mathrm{d}x$ où R est une fraction rationnelle à deux indéterminées, ou de la forme $\int R(\mathrm{e}^x)\,\mathrm{d}x$ où R est une fonction rationnelle, la méthode générale est simple : on fait le changement de variable $t = \mathrm{e}^x$, qui permet ici aussi de se ramener au calcul de l'intégrale d'une fraction rationnelle.

■ Comment calculer une intégrale ?

Cette problématique généralise la précédente ; mais calculer une intégrale, ce n'est pas forcément calculer une primitive. Dans certains cas, on ne sait pas calculer une primitive mais on sait calculer une intégrale entre certaines bornes particulières.

SF10.7 Utiliser directement une primitive

Il s'agit de la situation la plus simple. Mais beaucoup d'entre vous n'arrivent pas à repérer une primitive « évidente ». De plus, dans beaucoup de cas, il faut faire quelques calculs pour voir apparaître une primitive usuelle.

Exemple

Calculer $I = \displaystyle\int_{\frac{\pi}{4}}^{\frac{\pi}{3}} \frac{\cos^4 x}{\sin^6 x}\,\mathrm{d}x$.

La fonction $x \mapsto \dfrac{\cos^4 x}{\sin^6 x}$ est définie et continue sur $\left[\dfrac{\pi}{4}\,;\dfrac{\pi}{3}\right]$, donc I existe.
On remarque que l'on peut écrire $\dfrac{\cos^4 x}{\sin^6 x} = \cot^4 x \cdot \dfrac{1}{\sin^2 x}$ et puisque la dérivée de la fonction cot *est $x \mapsto -\dfrac{1}{\sin^2 x}$ on a directement une primitive :*

$$I = \left[-\frac{\cot^5 x}{5}\right]_{\frac{\pi}{4}}^{\frac{\pi}{3}} = -\frac{1}{5}\left(\left(\frac{1}{\sqrt{3}}\right)^5 - 1\right) = \frac{1}{5} - \frac{\sqrt{3}}{135}\cdot$$

Exemple

Calculer $I = \displaystyle\int_2^{+\infty} \frac{2x-4}{x^3-3x+2}\,\mathrm{d}x$.

- *Pour tout $x \in \mathbb{R}$ on a $x^3-3x+2 = (x-1)(x^2+x-2) = (x-1)^2(x+2)$. La fonction $f\colon x \mapsto \dfrac{2x-4}{x^3-3x+2}$ est donc continue sur $[2\,;+\infty[$. De plus, $f(x) \underset{x\to+\infty}{\sim} \dfrac{2}{x^2}$ et $x \mapsto \dfrac{1}{x^2}$ est intégrable au voisinage de $+\infty$, d'où l'existence de I.*
- *La décomposition en éléments simples de $f(x)$ s'écrit :*

$$f(x) = \frac{2x-4}{(x-1)^2(x+2)} = \frac{a}{x-1} + \frac{b}{(x-1)^2} + \frac{c}{x+2}$$

et l'on trouve aisément : $a = \dfrac{8}{9}$, $b = -\dfrac{2}{3}$ et $c = -\dfrac{8}{9}$. On a alors :

$$I = \frac{8}{9}\int_2^{+\infty}\left(\frac{1}{x-1}-\frac{1}{x+2}\right)\mathrm{d}x - \frac{2}{3}\int_2^{+\infty}\frac{1}{(x-1)^2}\,\mathrm{d}x$$
$$= \frac{8}{9}\left[\ln\frac{x-1}{x+2}\right]_2^{+\infty} + \frac{2}{3}\left[\frac{1}{x-1}\right]_2^{+\infty}.$$

✓ *Attention ici à ne pas « couper en deux » l'intégrale* $\int_2^{+\infty}\left(\frac{1}{x-1}-\frac{1}{x+2}\right)\mathrm{d}x$, *ce qui donnerait deux intégrales divergentes !*

Finalement, on trouve : $I = \frac{16}{9}\ln 2 - \frac{2}{3}$.

SF10.8 Déterminer une primitive

Dans la plupart des cas, il faut déjà transformer l'intégrale pour pouvoir la calculer. Pour cela, on utilise les deux techniques majeures du calcul intégral : l'intégration par parties et le changement de variable.

Il est absolument indispensable de bien maîtriser ces deux méthodes ; pour cela, il faut faire un nombre conséquent d'exercices afin, non seulement de maîtriser la technique calculatoire, mais aussi et surtout afin de repérer les situations où elles peuvent s'appliquer (outre celles déjà exposées au début de ce chapitre).

Nous vous invitons à chercher les exemples traités ci-dessous afin de vous entraîner, sans regarder la solution dans un premier temps.

Exemple

Calculer $I = \displaystyle\int_0^1 \ln(1+x^2)\,\mathrm{d}x$.

Une intégration par parties s'impose, et donne :

$$I = -\int_0^1 \frac{x\cdot 2x}{1+x^2}\,\mathrm{d}x + \Big[x\ln(1+x^2)\Big]_0^1.$$

Le terme entre crochets vaut $\ln 2$ *et l'intégrale vaut :*

$$-\int_0^1 \frac{2x^2}{1+x^2}\,\mathrm{d}x = -2\int_0^1 \frac{1+x^2}{1+x^2}\,\mathrm{d}x + 2\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = -2 + 2\frac{\pi}{4}.$$

✓ *Notez l'astuce classique consistant à faire apparaître le dénominateur* $1+x^2$ *au numérateur, en écrivant* $x^2 = (x^2+1)-1$.

Finalement, on a : $I = \frac{\pi}{2} - 2 + \ln 2$.

Exemple

Calculer $I = \displaystyle\int_0^1 \frac{x^2}{(x^2+1)^3}\,\mathrm{d}x$.

La fonction considérée ici est continue sur $[0\,;1]$, *donc l'intégrale proposée existe.*

La méthode classique consistant à décomposer la fraction rationnelle $\frac{X^2}{(X^2+1)^3}$ *est applicable si l'on mène bien les calculs, mais pénible...*

✓ *Beaucoup d'entre vous voudraient, dans ce type d'intégrale, faire le changement de variable* $t = x^2$. *Mais c'est impossible ! En effet il n'y a a pas de terme en* $\mathrm{d}t = 2x\,\mathrm{d}x$ *dans l'expression, et ce changement de variable conduirait donc à faire apparaître des radicaux...*

Il faut ici prendre le temps d'examiner la fonction à intégrer et remarquer qu'une partie de la fonction est une dérivée. Plus précisément, on écrit :

$$\frac{x^2}{(x^2+1)^3} = x \cdot \frac{x}{(x^2+1)^3},$$

et on remarque que l'expression $\dfrac{x}{(x^2+1)^3}$ *est de la forme* $\dfrac{u'}{u^3} = u'u^{-3}$, *à une constante près. Cela suggère de faire une intégration par parties :*

$$I = \int_0^1 \underbrace{x}_{=u} \cdot \underbrace{\frac{x}{(x^2+1)^3}}_{=v'}\,\mathrm{d}x = \left[-\frac{x}{4(x^2+1)^2}\right]_0^1 + \frac{1}{4}\int_0^1 \frac{\mathrm{d}x}{(x^2+1)^2}\cdot$$

Le terme entre crochets vaut $-\dfrac{1}{16}$ *; la seconde intégrale se calcule en faisant encore une intégration par parties, après avoir utilisé l'astuce classique consistant à écrire* $1 = (1+x^2) - x^2$ *:*

$$\begin{aligned}\int_0^1 \frac{\mathrm{d}x}{(x^2+1)^2} &= \int_0^1 \frac{\mathrm{d}x}{x^2+1} - \int_0^1 \underbrace{x}_{=u}\cdot\underbrace{\frac{x}{(x^2+1)^2}}_{=v'}\,\mathrm{d}x\\ &= \big[\operatorname{Arctan} x\big]_0^1 - \left\{\left[-\frac{x}{2(x^2+1)}\right]_0^1 + \frac{1}{2}\int_0^1 \frac{\mathrm{d}x}{x^2+1}\right\}\\ &= \frac{\pi}{4} + \frac{1}{4} - \frac{1}{2}\cdot\frac{\pi}{4} = \frac{\pi}{8} + \frac{1}{4}\cdot\end{aligned}$$

En rassemblant le tout, on trouve : $I = \dfrac{\pi}{32}\cdot$

Exemple

Calculer $I = \displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{\sqrt{1+\mathrm{e}^{2t}}}\cdot$

- *La fonction* $t \mapsto \dfrac{1}{\sqrt{1+\mathrm{e}^{2t}}}$ *est continue positive sur* $\mathbb{R}_+$, *et elle est équivalente lorsque* $t \to +\infty$ *à* $t \mapsto \mathrm{e}^{-t}$ *qui est une fonction de référence intégrable, d'où l'existence de* I.
- *En combinant directement les changements de variable vus en* **[SF10.4]** *et en***[SF10.6]**, *on peut poser* $u = \sqrt{1+\mathrm{e}^{2t}}$.

  On obtient : $I = \displaystyle\int_{\sqrt{2}}^{+\infty} \frac{\mathrm{d}u}{u^2-1} = \left[\frac{1}{2}\ln\frac{u-1}{u+1}\right]_{\sqrt{2}}^{+\infty} = \ln(\sqrt{2}+1).$

Exemple

Calculer $I = \displaystyle\int_{-\pi}^{\pi} \frac{\mathrm{d}t}{1+\cos\alpha\cos t}$ où $\alpha \in \left]0\,;\frac{\pi}{2}\right[$.

Compte tenu des hypothèses faites sur α, *la fonction* $t \mapsto \dfrac{1}{1+\cos\alpha\cos t}$ *est définie et continue sur* $[-\pi\,;\pi]$ *donc* I *existe.*

On fait le changement de variable classique $u = \tan\dfrac{t}{2}$, *qui réalise une bijection de classe* $\mathscr{C}^1$ *strictement monotone de* $]-\pi\,;\pi[$ *sur* $\mathbb{R}$ *(cette dernière remarque assurant l'existence de l'intégrale généralisée qui va suivre).*

On a alors $t = 2\operatorname{Arctan} u$ *d'où* $\mathrm{d}t = 2\dfrac{\mathrm{d}u}{1+u^2}$ *et* $\cos t = \dfrac{1-u^2}{1+u^2}$ *donc :*

$$I = 2\int_{-\infty}^{+\infty} \frac{\mathrm{d}u}{u^2(1-\cos\alpha)+1+\cos\alpha} = \frac{2}{1-\cos\alpha}\int_{-\infty}^{+\infty} \frac{\mathrm{d}u}{u^2+\frac{1+\cos\alpha}{1-\cos\alpha}}$$
$$= \frac{2}{1-\cos\alpha}\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}}\left[\operatorname{Arctan}\left(u\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}}\right)\right]_{-\infty}^{+\infty}.$$

On obtient donc $I = \dfrac{2\pi}{\sin\alpha}$.

Exemple

Calculer $I = \displaystyle\int_0^1 \frac{\mathrm{d}x}{x+\sqrt{1-x^2}}$.

L'existence de I *ne pose pas de problème puisque, pour tout* $x \in [0\,;1]$, $x+\sqrt{1-x^2} > 0$.
On applique la méthode citée en **[SF10.4]**, *en posant* $x = \sin t$. *On obtient :*

$$I = \int_0^{\frac{\pi}{2}} \frac{\cos t}{\sin t+\cos t}\,\mathrm{d}t\,.$$

On pourrait ici poser classiquement $u = \tan\dfrac{t}{2}$, *mais il est plus astucieux d'introduire l'intégrale « complémentaire »* $J = \displaystyle\int_0^{\frac{\pi}{2}} \frac{\sin t}{\sin t+\cos t}\,\mathrm{d}t$.
En effet, le changement de variable $t \mapsto \frac{\pi}{2} - t$ *montre que* $I = J$ *d'où*

$$2I = I + J = \int_0^{\frac{\pi}{2}} \frac{\sin t+\cos t}{\sin t+\cos t} = \frac{\pi}{2}$$

et finalement $I = \dfrac{\pi}{4}$.

Exemple

Calculer $I = \displaystyle\int_1^{+\infty} \frac{\mathrm{d}x}{x(x^2+1)^2}$.

- *La fonction* $x \mapsto \dfrac{1}{x(x^2+1)^2}$ *est continue positive sur* $[1\,;+\infty[$, *et elle est équivalente lorsque* $x \to +\infty$ *à la fonction* $x \mapsto \dfrac{1}{x^3}$, *qui est une fonction de Riemann positive intégrable au voisinage de* $+\infty$. *Cela prouve que* I *existe.*
- *Là encore, il ne faut pas se lancer d'emblée dans une décomposition en éléments simples, mais examiner l'intégrale proposée. Et avec un peu de réflexion, vous* devez *arriver à remarquer qu'il y a un changement de variable incontournable... Plus précisément, il faut penser à écrire :*

$$\frac{\mathrm{d}x}{x(x^2+1)^2} = \underbrace{\frac{1}{x^2(x^2+1)^2}}_{\text{fonction de } x^2} \cdot \underbrace{x\,\mathrm{d}x}_{=\frac{1}{2}\,\mathrm{d}(x^2)} \cdot$$

Sous cette forme, le changement de variable $t = x^2$ apparaît évident ! Il ne reste plus qu'à faire les calculs.

- *L'application $x \mapsto x^2$ étant une bijection strictement monotone de classe $\mathscr{C}^1$ de l'intervalle $[1\,;+\infty[$ sur lui-même, le changement de variable dans l'intégrale généralisée est licite, et l'on obtient :*

$$I = \frac{1}{2}\int_1^{+\infty} \frac{\mathrm{d}t}{t(t+1)^2} = \frac{1}{2}\int_1^{+\infty} \left(\frac{1}{t} - \frac{1}{t+1} - \frac{1}{(t+1)^2}\right)\mathrm{d}t\,.$$

✓ *Exceptionnellement, nous ne faisons pas tous les calculs ici, le but de cet exemple étant essentiellement de vous montrer comment on peut trouver un changement de variable intéressant.*

On trouve : $I = \dfrac{\ln 2}{2} - \dfrac{1}{4}\cdot$

SF10.9 Trouver une relation de récurrence

Il s'agit de la méthode à essayer d'emblée dès que l'intégrale à calculer dépend d'un entier n, sauf dans certains cas où un calcul direct est possible.
Dans la plupart des cas, pour établir une relation de récurrence, on utilise une intégration par parties.

Exemple

Calculer $I_{p,q} = \displaystyle\int_0^1 t^p(1-t)^q\,\mathrm{d}t$ pour $(p,q) \in \mathbb{N}^2$.

Pour $p \in \mathbb{N}^*$ et $q \in \mathbb{N}$ on effectue une intégration par parties en prenant $u(t) = t^p$ et $v'(t) = (1-t)^q$. On obtient :

$$I_{p,q} = \underbrace{\left[-\frac{1}{q+1}(1-t)^{q+1}t^p\right]_0^1}_{=0 \text{ car } p\geqslant 1} + \frac{1}{q+1}\int_0^1 pt^{p-1}(1-t)^{q+1}\,\mathrm{d}t\,,$$

d'où la relation de récurrence : $I_{p,q} = \dfrac{p}{q+1} I_{p-1,q+1}\cdot$

On en déduit :

$$I_{p,q} = \frac{p}{q+1}\cdot\frac{p-1}{q+2}\cdots\frac{1}{q+p} I_{0,q+p} = \frac{p!q!}{(p+q)!} I_{0,q+p}\,.$$

Or $I_{0,q+p} = \displaystyle\int_0^1 (1-t)^{p+q}\,\mathrm{d}t = \frac{1}{p+q+1}$ *donc :* $I_{p,q} = \dfrac{p!q!}{(p+q+1)!}\cdot$

SF10.10 Utiliser un changement de variable

Nous avons déjà vu l'intérêt de faire un changement de variable pour transformer une intégrale : simplifier l'expression à intégrer, changer les bornes, faire apparaître une quantité remarquable…
Dans le cas d'une intégrale définie, certains changements de variable astucieux permettent parfois de retrouver l'intégrale de départ, et donc d'obtenir une équation pour la calculer (voir aussi un exemple page 439).

Exemple

Existence et calcul de $I = \displaystyle\int_0^{\frac{\pi}{2}} \ln(\sin x)\,\mathrm{d}x$.

- *La fonction* $x \mapsto \ln(\sin x)$ *est continue sur* $\left]0\,;\frac{\pi}{2}\right]$ *; de plus* $\sin x \underset{x\to 0}{\sim} x$ *donc* $\ln(\sin x) \underset{x\to 0}{\sim} \ln x$ *; puisque la fonction* $x \mapsto \ln x$ *est intégrable et de signe constant sur* $\left]0\,;\frac{\pi}{2}\right]$*, cela prouve l'existence de l'intégrale proposée.*
- *Le changement de variable affine* $y = \dfrac{\pi}{2} - x$ *(qui conserve la nature de l'intégrale) donne :* $\displaystyle\int_0^{\frac{\pi}{2}} \ln(\sin x)\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \ln(\cos y)\,\mathrm{d}y$*, donc :*

$$\begin{aligned}
2I &= \int_0^{\frac{\pi}{2}} (\ln(\sin x) + \ln(\cos x))\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \ln(\sin x \cdot \cos x)\,\mathrm{d}x \\
&= \int_0^{\frac{\pi}{2}} \ln\left(\frac{\sin 2x}{2}\right)\mathrm{d}x = \int_0^{\frac{\pi}{2}} \ln(\sin 2x)\,\mathrm{d}x - \frac{\pi}{2}\ln 2 \\
&= \frac{1}{2}\int_0^{\pi} \ln(\sin t)\,\mathrm{d}t - \frac{\pi}{2}\ln 2 \quad \textit{(changement de variable licite } t = 2x\textit{)} \\
&= \frac{1}{2}\int_0^{\frac{\pi}{2}} \ln(\sin t)\,\mathrm{d}t + \frac{1}{2}\int_{\frac{\pi}{2}}^{\pi} \ln(\sin t)\,\mathrm{d}t - \frac{\pi}{2}\ln 2 \\
&= \frac{1}{2}I + \frac{1}{2}\underbrace{\int_0^{\frac{\pi}{2}} \ln(\sin u)\,\mathrm{d}u}_{=I} - \frac{\pi}{2}\ln 2 \quad \textit{(changement licite } u = \pi - t\textit{)}
\end{aligned}$$

d'où l'on tire : $I = -\dfrac{\pi}{2}\ln 2$.

SF10.11 Utiliser des sommes de Riemann

Nous avons déjà vu en **[SF8.5]** que l'on pouvait utiliser les intégrales pour calculer la limite de certaines suites ; il s'agit d'ailleurs de l'utilisation la plus courante des sommes de Riemann.

Cependant dans certains cas (assez rares, et une indication doit vous être donnée), on peut aussi calculer une intégrale en l'écrivant comme limite d'une suite de sommes de Riemann.

Exemple

Soit z un nombre complexe tel que $|z| \neq 1$.

Calculer $I = \displaystyle\int_0^{2\pi} \frac{\mathrm{d}t}{z - \mathrm{e}^{\mathrm{i}t}}$.

On remarque déjà que l'intégrale proposée a bien un sens puisque lorsque t *décrit* $[0\,;2\pi]$ *le dénominateur* $z - \mathrm{e}^{\mathrm{i}t}$ *ne s'annule pas compte tenu de l'hypothèse faite sur* z*.*

Soit $n \in \mathbb{N}^*$*. En considérant la subdivision de l'intervalle* $[0\,;2\pi]$ *formée des* n *points* $\dfrac{2k\pi}{n}$ *pour* $k \in [\![0\,;n-1]\!]$ *on obtient* $I = \lim\limits_{n\to+\infty} S_n$ *avec :*

$$S_n = \frac{2\pi}{n} \sum_{k=0}^{n-1} \frac{1}{z - \omega_k},$$

en ayant posé $\omega_k = \mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}$ *: ce sont les racines* n^{es} *de l'unité.*
Or si l'on considère le polynôme $P = \prod_{k=0}^{n-1} (X - \omega_k) \in \mathbb{C}[X]$*, il est facile de vérifier à l'aide de la formule donnant la dérivée d'un produit que :*

$$\frac{P'}{P} = \sum_{k=0}^{n-1} \frac{1}{X - \omega_k}.$$

Mais puisque $P = X^n - 1$, $\dfrac{P'}{P} = \dfrac{nX^{n-1}}{X^n - 1}$ *donc finalement :*

$$I = \lim_{n \to +\infty} 2\pi \frac{z^{n-1}}{z^n - 1} = \begin{cases} 0 & \text{si } |z| < 1 \\ \dfrac{2\pi}{z} & \text{si } |z| > 1. \end{cases}$$

SF10.12 Utiliser une intégrale dépendant d'un paramètre

Les intégrales dépendant d'un paramètre seront étudiées en détail en **[SF14.21]**. Nous donnons seulement ci-dessous un exemple où l'utilisation des techniques concernant ces intégrales peut servir à calculer une intégrale bien que leur utilisation n'apparaisse pas de prime abord.

Exemple

a et b étant des réels strictement supérieurs à 1, calculer $I = \displaystyle\int_0^\pi \ln\left(\frac{b - \cos t}{a - \cos t}\right) \mathrm{d}t$.

Compte tenu des hypothèses faites sur a *et* b*, la fonction* $t \mapsto \ln\left(\dfrac{b - \cos t}{a - \cos t}\right)$ *est continue sur* $[0\,;\pi]$*, ce qui assure l'existence de l'intégrale.*
On remarque que l'on peut écrire $I = F(b) - F(a)$ *où* $F(x) = \displaystyle\int_0^\pi \ln(x - \cos t)\,\mathrm{d}t$ *pour* $x > 1$.
Il ne reste plus qu'à calculer F*. À l'aide du théorème de dérivabilité d'une intégrale à paramètre rappelé en* **[S14.11]** *on peut montrer que* F *est de classe* $\mathscr{C}^1$ *sur* $]1\,;+\infty[$*; nous vous invitons à le vérifier complètement (appliquer le théorème sur tout segment inclus dans* $]1\,;+\infty[$*).*
Le même théorème donne aussi :

$$\forall\, x \in]1\,;+\infty[,\ F'(x) = \int_0^\pi \frac{\mathrm{d}t}{x - \cos t}.$$

Pour calculer cette intégrale on fait le changement de variable classique $u = \tan\dfrac{t}{2}$ *qui donne :*

$$F'(x) = \int_0^{+\infty} \frac{2\,\mathrm{d}u}{(1+u^2)\left(x - \frac{1-u^2}{1+u^2}\right)} = \int_0^{+\infty} \frac{2\,\mathrm{d}u}{u^2(x+1) + x - 1}$$

$$= \frac{2}{x+1} \int_0^{+\infty} \frac{\mathrm{d}u}{u^2 + \frac{x-1}{x+1}} = \frac{2}{x+1}\sqrt{\frac{x+1}{x-1}} \left[\operatorname{Arctan}\left(u\sqrt{\frac{x+1}{x-1}}\right)\right]_0^{+\infty}$$

soit $F'(x) = \dfrac{\pi}{\sqrt{x^2 - 1}}$.

Une primitive de $x \mapsto \dfrac{1}{\sqrt{x^2-1}}$ *sur* $]1\,;+\infty[$ *est* $x \mapsto \ln\left(x+\sqrt{x^2-1}\right)$ *donc*

$$I = F(b) - F(a) = \left[\pi \ln\left(x+\sqrt{x^2-1}\right)\right]_a^b = \pi \ln\left(\frac{b+\sqrt{b^2-1}}{a+\sqrt{a^2-1}}\right).$$

SF10.13 Utiliser une série

L'utilisation d'une série de fonctions pour calculer une intégrale est une méthode importante, à laquelle on ne pense pas toujours spontanément.
Nous verrons en détail l'utilisation de séries entières en **[SF13.18]** et d'autres séries de fonctions en **[SF14.22]**.

Nous donnons ci-dessous un autre exemple de calcul d'une intégrale à l'aide d'une série. Cet exemple n'utilise pas les séries de fonctions, mais la relation de Chasles.

Exemple

Calculer $I = \displaystyle\int_0^1 \frac{\mathrm{d}t}{2+\left\lfloor \frac{1}{t} \right\rfloor}$.

Notons déjà que la fonction à intégrer est continue par morceaux sur $]0\,;1]$ *(c'est même une fonction en escalier), et se prolonge en* 0 *puisque* $\lim\limits_{t\to 0^+} \left\lfloor \frac{1}{t} \right\rfloor = +\infty$ *(par le théorème d'encadrement) donc* $\lim\limits_{t\to 0^+} \dfrac{1}{2+\left\lfloor \frac{1}{t} \right\rfloor} = 0$.

L'intégrale étant une fonction continue de ses bornes, on a :

$$I = \lim_{n\to+\infty} \int_{1/n}^1 \frac{\mathrm{d}t}{2+\left\lfloor \frac{1}{t} \right\rfloor}.$$

Or d'après la relation de Chasles, pour tout $n \geqslant 2$ *:*

$$\int_{1/n}^1 \frac{\mathrm{d}t}{2+\left\lfloor \frac{1}{t} \right\rfloor} = \sum_{k=2}^{n} \int_{1/k}^{1/(k-1)} \frac{\mathrm{d}t}{2+\left\lfloor \frac{1}{t} \right\rfloor} = \sum_{k=2}^{n} \int_{1/k}^{1/(k-1)} \frac{\mathrm{d}t}{2+(k-1)},$$

donc $I = \lim\limits_{n\to+\infty} \sum\limits_{k=2}^{n} \dfrac{1}{k+1}\left(\dfrac{1}{k-1}-\dfrac{1}{k}\right) = \lim\limits_{n\to+\infty} \sum\limits_{k=2}^{n} \dfrac{1}{(k-1)k(k+1)}$.
Une décomposition en éléments simples donne :

$$\forall\, k \geqslant 2,\ \frac{1}{(k-1)k(k+1)} = \frac{1/2}{k-1} - \frac{1}{k} + \frac{1/2}{k+1},$$

d'où :

$$\begin{aligned}\sum_{k=2}^{n} \frac{1}{(k-1)k(k+1)} &= \frac{1}{2}\sum_{k=1}^{n-1} \frac{1}{k} - \sum_{k=2}^{n} \frac{1}{k} + \frac{1}{2}\sum_{k=3}^{n+1} \frac{1}{k} \\ &= \frac{1}{2}\left(1+\frac{1}{2}\right) - \left(\frac{1}{2}+\frac{1}{n}\right) + \frac{1}{2}\left(\frac{1}{n}+\frac{1}{n+1}\right).\end{aligned}$$

On obtient donc : $I = \dfrac{1}{4}$.

■ Comment aborder un exercice moins classique mettant en jeu des intégrales ?

SF10.14 Transformer l'intégrale

Dans de nombreux cas, il est possible, voire indispensable, de simplifier ou de modifier les intégrales proposées avant d'appliquer toute autre méthode. Cela peut se faire à l'aide d'un changement de variable ou d'une intégration par parties.

> ✓ Il s'agit d'une règle fondamentale : si vous êtes dans une impasse lors du calcul ou de l'étude d'une intégrale (généralisée ou non), il faut absolument penser aux deux techniques de base que sont l'intégration par parties et le changement de variable.

Nous ne donnons ici pour l'instant qu'un exemple de cette méthode, mais elle sera régulièrement utilisée par la suite.

Exemple

Soit f une application de classe $\mathscr{C}^1$ de $[a\,;b]$ dans $\mathbb{C}$.

Démontrer que : $\lim\limits_{n\to+\infty} \displaystyle\int_a^b f(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t = 0$ (il s'agit du *lemme de Lebesgue*).

Puisque l'énoncé met en avant la dérivabilité de f, il est naturel de penser à faire une intégration par parties. On obtient, pour $n \in \mathbb{N}^*$:

$$\int_a^b f(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t = \left[f(t)\frac{\mathrm{e}^{\mathrm{i}nt}}{\mathrm{i}n}\right]_a^b - \frac{1}{\mathrm{i}n}\int_a^b f'(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t\,,$$

d'où :

$$\begin{aligned}\left|\int_a^b f(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t\right| &\leqslant \left|\frac{f(b)\mathrm{e}^{\mathrm{i}nb} - f(a)\mathrm{e}^{\mathrm{i}na}}{\mathrm{i}n}\right| + \frac{1}{n}\left|\int_a^b f'(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t\right| \\ &\leqslant \frac{|f(b)|+|f(a)|}{n} + \frac{1}{n}\int_a^b |f'(t)|\underbrace{\left|\mathrm{e}^{\mathrm{i}nt}\right|}_{=1}\mathrm{d}t \\ &\leqslant \frac{|f(b)|+|f(a)|}{n} + \frac{b-a}{n}\,\|f'\|_\infty^{[a;b]}\end{aligned}$$

($\|f'\|_\infty^{[a;b]}$ existe puisque f' est continue sur le segment $[a\,;b]$).

Cette inégalité permet de conclure : $\lim\limits_{n\to+\infty} \displaystyle\int_a^b f(t)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t = 0$.

SF10.15 Utiliser des inégalités

Le sujet est vaste, l'obtention d'inégalités judicieuses étant une des bases de l'analyse. On peut penser à utiliser la positivité de l'intégrale, l'inégalité de la moyenne, ou l'inégalité de Cauchy-Schwarz. Dans certains cas, on peut aussi être amené à découper l'intervalle pour appliquer ces inégalités sur chacun des sous-intervalles.

Exemple

Pour $n \in \mathbb{N}^*$ soit $I_n = \displaystyle\int_0^{+\infty} \frac{\sin(nx)}{nx+x^2}\,\mathrm{d}x$.

Justifier l'existence de I_n et calculer $\lim\limits_{n\to+\infty} I_n$.

- *L'existence de cette intégrale ne doit pas poser trop de problème, car les méthodes utilisées sont classiques (voir* **[SF14.3]***). En effet :*
  - *la fonction $x \mapsto \dfrac{\sin(nx)}{nx+x^2}$ est continue sur $]0\,;+\infty[$;*
  - *$\dfrac{\sin(nx)}{nx} \underset{x\to 0}{\sim} \dfrac{nx}{nx+x^2}$ donc $\lim\limits_{x\to 0^+} \dfrac{\sin(nx)}{nx+x^2} = 1$, et la fonction est prolongeable par continuité en 0^+ ;*
  - *$\forall\, x > 0$, $\left|\dfrac{\sin(nx)}{nx+x^2}\right| \leqslant \dfrac{1}{x^2}$, et $x \mapsto \dfrac{1}{x^2}$ est positive et intégrable au voisinage de $+\infty$, donc l'intégrale proposée est absolument convergente.*
- *Nous allons faire des majorations différentes selon que l'on est au voisinage de 0 ou non.*
  - *En utilisant l'inégalité connue $|\sin t| \leqslant |t|$:*
$$\left|\int_0^{1/n} \frac{\sin(nx)}{nx+x^2}\,\mathrm{d}x\right| \leqslant \int_0^{1/n} \left|\frac{\sin nx}{nx}\right| \mathrm{d}x \leqslant \int_0^{1/n} \mathrm{d}x = \frac{1}{n}\,,$$
  donc $\lim\limits_{n\to+\infty} \left|\displaystyle\int_0^{1/n} \frac{\sin(nx)}{nx+x^2}\,\mathrm{d}x\right| = 0$.
  - *Puis :*
$$\left|\int_{1/n}^{+\infty} \frac{\sin(nx)}{nx+x^2}\,\mathrm{d}x\right| \leqslant \int_{1/n}^{+\infty} \frac{1}{nx+x^2}\,\mathrm{d}x = \int_{1/n}^{+\infty} \frac{1}{n}\left(\frac{1}{x} - \frac{1}{n+x}\right)\mathrm{d}x$$
$$\leqslant \frac{1}{n}\left[\ln\frac{x}{n+x}\right]_{1/n}^{+\infty} = \frac{\ln(n^2+1)}{n}\,,$$
  donc on a aussi $\lim\limits_{n\to+\infty} \left|\displaystyle\int_{1/n}^{+\infty} \frac{\sin(nx)}{nx+x^2}\,\mathrm{d}x\right| = 0$, ce qui permet de conclure : $\lim\limits_{n\to+\infty} I_n = 0$.

Exemple

Déterminer les fonctions $f \in \mathscr{C}([0\,;1],\mathbb{R})$ telles que $\displaystyle\int_0^1 f(x)\,\mathrm{d}x = \frac{1}{3} + \int_0^1 [f(x^2)]^2\,\mathrm{d}x$.

L'inégalité de Cauchy-Schwarz appliquée aux fonctions $x \mapsto x$ et $x \mapsto f(x^2)$ donne :

$$\left(\int_0^1 x f(x^2)\,\mathrm{d}x\right)^2 \leqslant \left(\int_0^1 x^2\,\mathrm{d}x\right)\left(\int_0^1 [f(x^2)]^2\,\mathrm{d}x\right) = \frac{1}{3}\int_0^1 [f(x^2)]^2\,\mathrm{d}x \quad (*)$$

En notant $I = \displaystyle\int_0^1 f$, le changement de variable $u = x^2$ donne $\displaystyle\int_0^1 x f(x^2)\,\mathrm{d}x = \frac{1}{2}I$.

Donc la relation de l'énoncé : $I = \frac{1}{3} + \int_0^1 [f(x^2)]^2 \,\mathrm{d}x$ *ainsi que l'inégalité* $(*)$ *conduisent à :*

$$\frac{1}{4}I^2 \leqslant \frac{1}{3}\left(I - \frac{1}{3}\right) \quad \textit{soit} \quad \left(I - \frac{2}{3}\right)^2 \leqslant 0\,.$$

On a donc forcément l'égalité $I = \frac{2}{3}$ *et par suite il y a aussi égalité dans l'inégalité de Cauchy-Schwarz initiale. Les fonctions concernées sont donc proportionnelles, c'est-à-dire qu'il existe un réel* λ *tel que pour tout* $x \in [0\,;1]$ *on ait :* $f(x^2) = \lambda x$ *soit aussi* $f(x) = \lambda\sqrt{x}$.

Réciproquement, en remplaçant f *par cette expression dans l'équation proposée, on trouve que seule la valeur* $\lambda = 1$ *convient.*

En conclusion, seule la fonction $f\colon x \mapsto \sqrt{x}$ *vérifie l'égalité proposée.*

SF10.16 Utiliser la relation de Chasles

Bien sûr, vous connaissez la relation de Chasles. Nous l'avons utilisée dans un exemple de **[SF10.13]** page 443 : son emploi était presque obligatoire du fait de la présence de la fonction partie entière. Mais il existe de nombreux exemples où il faut penser à l'utiliser bien que cela n'apparaisse pas directement dans l'énoncé (voyez le 1er exemple de **[SF10.15]** page 445).

Exemple

1) Calculer $J = \displaystyle\int_0^{\pi} \frac{\mathrm{d}x}{1+\cos^2 x}$.

2) Soit $f \in \mathscr{C}^1([0\,;\pi],\mathbb{R})$. Calculer $\displaystyle\lim_{n\to+\infty} \int_0^{\pi} \frac{f(x)}{1+\cos^2 nx}\,\mathrm{d}x$.

1) *L'existence de* J *ne pose pas de problème.*
Le changement de variable $x \mapsto \pi - x$ *montre que l'intégrale de* 0 *à* $\frac{\pi}{2}$ *est égale à celle de* $\frac{\pi}{2}$ *à* π *donc* $J = 2\displaystyle\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}x}{1+\cos^2 x}$.
Pour le calcul, en utilisant les règles de Bioche, on est amené à poser $t = \tan x$, *qui est un changement de variable de classe* $\mathscr{C}^1$ *et strictement monotone de* $\left[0\,;\frac{\pi}{2}\right[$ *sur* $\mathbb{R}_+$ *d'où :*

$$J = 2\int_0^{+\infty} \frac{\mathrm{d}t}{(1+t^2)\left(1+\frac{1}{1+t^2}\right)} = 2\int_0^{+\infty} \frac{\mathrm{d}t}{2+t^2}$$

$$= \frac{2}{\sqrt{2}}\left[\operatorname{Arctan}\left(\frac{t}{\sqrt{2}}\right)\right]_0^{+\infty} = \frac{\pi}{\sqrt{2}}\,.$$

2) *Le changement de variable* $t = nx$ *donne (en notant* I_n *l'intégrale proposée) :* $I_n = \frac{1}{n}\displaystyle\int_0^{n\pi} \frac{f\left(\frac{t}{n}\right)}{1+\cos^2 t}\,\mathrm{d}t$, *donc par la relation de Chasles :* $I_n = \frac{1}{n}\displaystyle\sum_{k=0}^{n-1} u_k$ *avec :*

$$u_k = \int_{k\pi}^{(k+1)\pi} \frac{f\left(\frac{t}{n}\right)}{1+\cos^2 t}\,\mathrm{d}t = \int_0^{\pi} \frac{f\left(\frac{u+k\pi}{n}\right)}{1+\cos^2 u}\,\mathrm{d}u\,.$$

Comme cela ressemble beaucoup à une somme de Riemann, on introduit $v_k = \displaystyle\int_0^\pi \frac{f\left(\frac{k\pi}{n}\right)}{1+\cos^2 u}\,\mathrm{d}u = f\left(\frac{k\pi}{n}\right) J$.

On a alors, en utilisant l'inégalité des accroissements finis :

$$\begin{aligned}|u_k - v_k| &\leqslant \int_0^\pi \frac{\left|f\left(\frac{u+k\pi}{n}\right) - f\left(\frac{k\pi}{n}\right)\right|}{1+\cos^2 u}\,\mathrm{d}u \\ &\leqslant \|f'\|_\infty \int_0^\pi \frac{\frac{u}{n}}{1+\cos^2 u}\,\mathrm{d}u \leqslant \frac{\pi\,\|f'\|_\infty}{n} J\,.\end{aligned}$$

On aura ensuite :

$$\left|I_n - \frac{1}{n}\sum_{k=0}^{n-1} v_k\right| \leqslant \frac{1}{n}\sum_{k=0}^{n-1} |u_k - v_k| \leqslant \frac{\pi\,\|f'\|_\infty}{n} J\,,$$

de sorte que $\displaystyle\lim_{n\to+\infty} I_n = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} v_k$ *(sous réserve d'existence). Mais*

$$\frac{1}{n}\sum_{k=0}^{n-1} v_k = \left(\frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{k\pi}{n}\right)\right) J$$

est au facteur π *près une somme de Riemann associée à* f *sur* $[0\,;\pi]$ *et finalement :*

$$\lim_{n\to+\infty} I_n = \left(\frac{1}{\pi}\int_0^\pi f\right) J = \frac{1}{\sqrt{2}}\int_0^\pi f\,.$$

Exemple

Soit f une fonction continue de $\mathbb{R}$ dans $\mathbb{R}$; on suppose que f admet en $-\infty$ une limite ℓ et en $+\infty$ une limite ℓ'.

Calculer : $\displaystyle\lim_{\substack{x\to-\infty\\ y\to+\infty}} \int_x^y (f(t+1) - f(t-1))\,\mathrm{d}t$.

Notons $F(x,y) = \displaystyle\int_x^y (f(t+1) - f(t-1))\,\mathrm{d}t$.

- *On a déjà, à l'aide de changements de variables affines :*

$$\begin{aligned}F(x,y) &= \int_x^y f(t+1)\,\mathrm{d}t - \int_x^y f(t-1)\,\mathrm{d}t \\ &= \int_{x+1}^{y+1} f(t)\,\mathrm{d}t - \int_{x-1}^{y-1} f(t)\,\mathrm{d}t\,,\end{aligned}$$

puis à l'aide de la relation de Chasles :

$$\begin{aligned}F(x,y) &= \int_{x+1}^{x-1} f(t)\,\mathrm{d}t + \int_{x-1}^{y+1} f(t)\,\mathrm{d}t - \int_{x-1}^{y+1} f(t)\,\mathrm{d}t - \int_{y+1}^{y-1} f(t)\,\mathrm{d}t \\ &= \int_{y-1}^{y+1} f(t)\,\mathrm{d}t - \int_{x-1}^{x+1} f(t)\,\mathrm{d}t\,.\end{aligned}$$

- *En notant* φ *une primitive de* f *(qui existe car* f *est continue), on a :*

$$F(x,y) = \big(\varphi(y+1) - \varphi(y-1)\big) - \big(\varphi(x+1) - \varphi(x-1)\big).$$

Or d'après le théorème des accroissements finis, il existe $c_y \in]y-1\,;y+1[$ *tel que :*

$$\varphi(y+1)-\varphi(y-1)=\big((y+1)-(y-1)\big)\varphi'(c_y)=2f(c_y)\,,$$

et puisque $c_y \in]y-1\,;y+1[$, *on a* $\lim\limits_{y\to+\infty} c_y = +\infty$ *puis* $\lim\limits_{y\to+\infty} 2f(c_y)=2\ell'$, *d'où finalement* $\lim\limits_{y\to+\infty}\big(\varphi(y+1)-\varphi(y-1)\big)=2\ell'$.
On montre de même que : $\lim\limits_{x\to-\infty}\big(\varphi(x+1)-\varphi(x-1)\big)=2\ell$, *d'où*

$$\lim_{\substack{x\to-\infty\\ y\to+\infty}} F(x,y)=2(\ell'-\ell)\,.$$

En d'autres termes, l'intégrale généralisée $\displaystyle\int_{-\infty}^{+\infty}(f(t+1)-f(t-1))\,\mathrm{d}t$ *converge et vaut* $2(\ell'-\ell)$.

SF10.17 Considérer l'intégrale fonction de ses bornes

Il s'agit ici d'utiliser les résultats rappelés en **[S10.9]**.

Exemple

Soit f une fonction continue sur $[0\,;1]$ à valeurs réelles positives. On suppose qu'il existe un réel k tel que, pour tout $x\in[0\,;1]$ on ait : $f(x)\leqslant k\displaystyle\int_0^x f(t)\,\mathrm{d}t$.
Montrer que f est la fonction nulle.

En, posant $F(x)=\displaystyle\int_0^x f(t)\,\mathrm{d}t$, F *est une fonction de classe* $\mathscr{C}^1$ *sur* $[0\,;1]$ *qui vérifie* $F'\leqslant kF$, *ou* $F'-kF\leqslant 0$.
Par analogie avec les équations différentielles, on pense alors à considérer la fonction $g\colon x\mapsto F(x)\mathrm{e}^{-kx}$. *On a alors* $g'\leqslant 0$; g *est donc décroissante sur* $[0\,;1]$; *puisque* $g(0)=0$, g *est négative.*
Mais f *étant supposée positive, il en est de même de* F *donc de* g. *Par suite* g *est nulle, donc* F *aussi puis* $f=F'$ *aussi.*

Exemple

Déterminer les fonctions f continues de $\mathbb{R}$ dans $\mathbb{R}$ qui vérifient :

$$\forall\, x\in\mathbb{R},\ \forall\, a>0,\ f(x)=\frac{1}{2a}\int_{x-a}^{x+a} f(t)\,\mathrm{d}t\,.$$

On procède bien sûr par « analyse-synthèse ».

- *Soit* f *une fonction qui convient.* f *étant continue possède une primitive* F ; *la relation de l'énoncé s'écrit alors :*

$$\forall\, x\in\mathbb{R},\ \forall\, a>0,\ 2af(x)=F(x+a)-F(x-a)\,.$$

Cette égalité prouve d'abord que f *est forcément de classe* $\mathscr{C}^1$. *De plus, en la dérivant par rapport à* a *on obtient :*

$$\forall\, x\in\mathbb{R},\ \forall\, a>0,\ 2f(x)=f(x+a)+f(x-a)$$

puis en dérivant cette nouvelle égalité par rapport à a :

$$\forall\, x\in\mathbb{R},\ \forall\, a>0,\ f'(x+a)=f'(x-a)\,.$$

On en déduit que f' *est constante, donc que* f *est une fonction affine.*

- *Réciproquement, il est facile de vérifier directement que toutes les fonctions affines $x \mapsto \alpha x + \beta$ vérifient bien la relation proposée.*

■ Comment utiliser les formules de Taylor ?

Parmi ces trois formules, celle avec reste intégrale est la plus précise. C'est donc celle à privilégier, mais dans beaucoup de cas l'inégalité de Taylor-Lagrange, plus simple, peut suffire.
Quant à la formule de Taylor-Young, elle ne doit servir *que* pour les développements limités, c'est-à-dire lorsqu'il s'agit d'étudier le comportement d'une fonction au voisinage d'un point : c'est une formule locale, alors que les deux premières sont des formules globales.

SF10.18 Utiliser la formule de Taylor avec reste intégrale

Exemple

Établir que pour tout réel $x \in \left[0\,;\frac{\pi}{2}\right]$: $x - \dfrac{x^3}{6} \leqslant \sin x \leqslant x - \dfrac{x^3}{6} + \dfrac{x^5}{120}$.

La fonction sin *étant de classe $\mathscr{C}^\infty$, on peut écrire la formule de Taylor avec reste intégrale entre* 0 *et* x *:*

$$\sin x = x - \frac{x^3}{6} + \int_0^x \frac{(x-t)^4}{4!}\cos t\,\mathrm{d}t\,.$$

Or pour $t \in \left[0\,;\frac{\pi}{2}\right]$, $0 \leqslant \cos t \leqslant 1$ *donc pour* $x \in \left[0\,;\frac{\pi}{2}\right]$ *:*

$$0 \leqslant \int_0^x \frac{(x-t)^4}{4!}\cos t\,\mathrm{d}t \leqslant \int_0^x \frac{(x-t)^4}{4!}\,\mathrm{d}t = \frac{x^5}{120}\,,$$

ce qui donne la double inégalité de l'énoncé.

✓ *On pouvait bien sûr démontrer ces inégalités en étudiant les fonctions différences, mais il eût été dommage de ne pas utiliser une formule de Taylor puisqu'elle apparaît ici de façon naturelle dans l'énoncé.*

SF10.19 Utiliser l'inégalité de Taylor-Lagrange

Exemple

Soit $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^\infty$. On suppose qu'il existe un polynôme $P \in \mathbb{R}[X]$, de degré impair, tel que $\forall\, n \in \mathbb{N},\ \forall\, x \in \mathbb{R},\ \left|f^{(n)}(x)\right| \leqslant |P(x)|$.

Montrer que f est la fonction nulle.

P étant de degré impair possède une racine réelle a. Compte tenu de l'hypothèse de l'énoncé, on a alors $f^{(n)}(a) = 0$ pour tout entier n.

Soit $x \in \mathbb{R}$. En appliquant l'inégalité de Taylor-Lagrange à f sur l'intervalle $I = [a\,;x]$ (ou $[x\,;a]$), on a :

$$|f(x)| \leqslant \frac{|x-a|^n}{n!}\sup_{t\in I}\left|f^{(n)}(t)\right| \leqslant \frac{|x-a|^n}{n!}\sup_{t\in I}|P(t)|\,.$$

Puisque $\sup_{t\in I}|P(t)|$ *ne dépend pas de* n*, il suffit de faire tendre* n *vers* $+\infty$ *dans l'inégalité précédente pour obtenir :* $f(x)=0$.

SF10.20 Utiliser la formule de Taylor-Young

Exemple

Soit $f:\mathbb{R}_+\to\mathbb{R}_+$ de classe $\mathscr{C}^\infty$ sur $\mathbb{R}_+$. On suppose que :

$$\forall\, x>0,\ f(x)>0 \quad \text{et} \quad \lim_{x\to 0^+} x^2\ln\big(f(x)\big)=-1\,.$$

Montrer que : $\forall\, p\in\mathbb{N},\ f^{(p)}(0)=0$.

Procédons par récurrence sur p.

- f *est continue en* 0*, donc* $\lim\limits_{x\to 0^+} f(x)=f(0)\geqslant 0$*. Si* $f(0)$ *était strictement positif, on aurait* $\lim\limits_{x\to 0^+}\ln\big(f(x)\big)=\ln\big(f(0)\big)$ *donc* $\lim\limits_{x\to 0^+} x^2\ln\big(f(x)\big)=0$*, c'est impossible.*
  On a donc $f(0)=0$*, ce qui démontre le résultat à l'ordre* $p=0$.
- *Supposons démontré* $f(0)=f'(0)=\ldots=f^{(p)}(0)=0$.
  f *étant de classe* $\mathscr{C}^\infty$*, la formule de Taylor-Young donne alors, au voisinage de* 0 *:*

$$f(x)=\frac{x^{p+1}}{(p+1)!}f^{(p+1)}(0)\big(1+\varphi(x)\big) \quad \text{avec} \quad \lim_{x\to 0^+}\varphi(x)=0\,.$$

  Si $f^{(p+1)}(0)$ *n'était pas nul, on aurait* $f^{(p+1)}(0)>0$ *et*

$$x^2\ln f^{(p+1)}(0)=x^2\ln(p+1)!+x^2\ln f(x)-(p+1)x^2\ln x-x^2\ln(1+\varphi(x))$$

  donc $\lim\limits_{x\to 0^+} x^2\ln f^{(p+1)}(0)=-1$*, c'est impossible.*
  On en déduit $f^{(p+1)}(0)=0$*, ce qui est la propriété à l'ordre* $p+1$*, et achève la récurrence.*

À vous de jouer !

Exercice 10.1 **SF10.1**

Étudier et tracer la courbe paramétrée par :

$$\begin{cases} x(t)=\cos 3t \\ y(t)=\sin t+\cos t \end{cases} \qquad \text{pour } t\in\mathbb{R}\,.$$

Préciser les coordonnées des points multiples.

Exercice 10.2 **SF10.1**

1) Étudier et tracer la courbe paramétrée par :

$$x(t) = \cos^2 t + \ln|\sin t| \quad \text{et} \quad y(t) = \sin t \cos t\,.$$

2) On note T (respectivement N) le point d'intersection avec l'axe des abscisses de la tangente (respectivement de la normale) en un point M de la courbe.

Montrer que la longueur du segment $[TN]$ est constante.

Exercice 10.3 **SF10.3** **SF10.4** **SF10.5** **SF10.8**

Calculer les intégrales suivantes.

$$\textbf{1)}\int_0^{\frac{\pi}{2}} \frac{\sin t}{1+\sin t+\cos t} \qquad \textbf{2)}\int_0^{\pi} \frac{\mathrm{d}t}{1+a^2\sin^2 t} \qquad \textbf{3)}\int_1^{+\infty} \frac{\sqrt{t^3+1}}{t^4}\,\mathrm{d}t$$

$$\textbf{4)}\int_1^{+\infty} \frac{t\ln t}{(1+t^2)^2}\,\mathrm{d}t$$

Exercice 10.4 **SF10.5** **SF10.9**

On pose, pour tout $n \in \mathbb{N}$: $a_n = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} \frac{\cos nx}{4\cos x - 5}\,\mathrm{d}x$.

1) Calculer a_0 et a_1.

2) Calculer $a_n + a_{n+2}$. En déduire une relation de récurrence sur les a_n.

3) Déterminer a_n pour tout $n \in \mathbb{N}$.

Exercice 10.5 **SF8.5** **SF10.11**

Pour $n \in \mathbb{N}^*$ on pose $u_n = \displaystyle\prod_{k=1}^{n}\left(1+\frac{\sqrt{k(n-k)}}{n^2}\right)$.

Prouver la convergence de la suite (u_n) et préciser sa limite.

Exercice 10.6 **SF10.5** **SF10.15** **SF10.16**

Pour $x > 0$, on pose : $f(x) = \displaystyle\int_0^{\frac{1}{x}} \frac{\mathrm{d}t}{x+\sin^2 t}$.

1) Montrer que f est bien définie et étudier sa monotonie.

2) Préciser les limites de f en 0^+ et en $+\infty$.

3) Déterminer un équivalent de f en $+\infty$ et en 0^+ (*on pourra poser* $\theta = \tan t$).

Exercice 10.7 **SF10.17** **SF13.20**

Quelles sont les fonctions f continues sur $\mathbb{R}$ telles que :

$$\forall\, x \in \mathbb{R},\ f(x) = -1 - \int_0^x (2x - t) f(t)\, \mathrm{d}t \ ?$$

Exercice 10.8 **SF10.18**

Soit f une application de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$, à valeurs dans $\mathbb{R}_+$.
On suppose qu'il existe $x_0 \in \mathbb{R}$ tel que $f(x_0) = 0$.

1) Que peut-on dire de $f'(x_0)$? de $f''(x_0)$?

2) On note : $m = \min\left\{k \in \mathbb{N}^* \;\middle|\; f^{(k)}(x_0) \neq 0\right\}$ (en supposant qu'il existe).
Montrer qu'il existe une fonction $g\colon \mathbb{R} \to \mathbb{R}$ de classe $\mathscr{C}^\infty$ telle que $f(x_0 + h) = h^m g(h)$. En déduire que :

$$\sqrt{f} \text{ est de classse } \mathscr{C}^\infty \text{ au voisinage de } x_0 \iff m \equiv 0 \ (\mathrm{mod}\ 4)\,.$$

Exercice 10.9 **SF10.19**

On considère une application de classe $\mathscr{C}^2$ de $\mathbb{R}$ dans $\mathbb{R}$ telle que f et f'' soient bornées.
On pose alors $M_0 = \sup_{x\in\mathbb{R}} |f(x)|$ et $M_2 = \sup_{x\in\mathbb{R}} |f''(x)|$.
En appliquant une formule de Taylor entre x et $x+h$ puis entre x et $x-h$, établir :

$$\forall\, x \in \mathbb{R},\ \forall\, h > 0,\ |f'(x)| \leqslant \frac{M_0}{h} + \frac{h}{2} M_2\,.$$

En déduire que f' est bornée sur $\mathbb{R}$ et que si l'on pose $M_1 = \sup_{x\in\mathbb{R}} |f'(x)|$, on a :

$$M_1 \leqslant \sqrt{2 M_0 M_2}\,.$$

Exercice 10.10 **SF9.2** **SF10.20**

1) Soit $n \in \mathbb{N}^*$. Soit I un intervalle centré en 0, et f une fonction de classe $\mathscr{C}^n$ de I dans $\mathbb{R}$ telle que : $f(0) = f'(0) = \ldots = f^{(n)}(0) = 0$.

On considère la fonction φ définie sur I par $\varphi(x) = \begin{cases} \dfrac{f(x)}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0\,. \end{cases}$

Montrer que φ est de classe $\mathscr{C}^{n-1}$ sur I.

2) Soit g de classe $\mathscr{C}^n$ sur I. Soit alors $\psi\colon x \in I \mapsto \begin{cases} \dfrac{g(x) - g(0)}{x} & \text{si } x \neq 0 \\ g'(0) & \text{si } x = 0\,. \end{cases}$

Montrer que ψ est de classe $\mathscr{C}^{n-1}$ sur I.

Thème 11 - Séries et familles sommables

■ Comment étudier la nature d'une série numérique ?

Cette question donne lieu à de nombreuses méthodes, et les situations sont extrêmement variées. Mais il s'agit d'une étude incontournable et il faut donc acquérir de bons réflexes et beaucoup de technicité. Pour cela, la résolution d'un nombre conséquent d'exercices différents est indispensable.
S'il est difficile de dire a priori quelle est la stratégie la plus efficace, on peut essayer de penser aux méthodes suivantes, en général dans cet ordre :

- regarder ce que fait le terme général ;
- regarder le signe du terme général ; si ce signe est constant à partir d'un certain rang, essayer d'utiliser une règle de comparaison ;
- penser à la convergence absolue, ce qui nous ramène à la situation ci-dessus ;
- obtenir un équivalent du terme général, ou plus généralement en faire un développement limité ;
- penser au critère spécial des séries alternées ;
- penser à la comparaison série/intégrale ;
- revenir enfin à la définition, en étudiant la suite des sommes partielles.

Très souvent, il faut savoir mêler plusieurs des méthodes ci-dessus : par exemple, on peut d'abord chercher un équivalent, puis appliquer l'une des autres méthodes à la série plus simple ainsi obtenue.

SF11.1 Repérer immédiatement une divergence grossière

On sait que si une suite (u_n) ne converge pas vers 0 alors la série $\sum u_n$ diverge. Donc on commence par regarder le comportement de u_n lorsque n tend vers $+\infty$.

✓ Il ne faut pas passer à côté de ce genre de situation.
Cependant, cela n'est à faire que si l'étude de la limite est rapide ! Si cela n'est pas immédiat, on passe directement à l'une des méthodes suivantes.

Exemple

Étudier la nature de la série $\sum_{n\geqslant 0} n!x^n$ selon les valeurs de $x \in \mathbb{R}$.

Si $x = 0$, la série est évidemment convergente.
Sinon, $\lim_{n\to+\infty} n!\,|x^n| = +\infty$ d'après les résultats de 1re année sur les croissances comparées des suites usuelles.
Ainsi, si $x \neq 0$, le terme général de la série ne tend pas vers zéro, donc cette série diverge.

Exemple

Étudier la nature de la série de terme général $\sin n$.

Nous allons démontrer que la suite de terme général $(\sin n)_{n\in\mathbb{N}}$ *n'a pas de limite. Cela entraînera donc la divergence grossière de la série* $\sum \sin n$.

Pour cela, procédons par l'absurde en supposant qu'il existe $\ell \in \mathbb{R}$ *tel que* $\lim\limits_{n\to+\infty} \sin n = \ell$.

Puisque $\sin(n+1) - \sin(n-1) = 2\sin 1 \cos n$ *(formules de trigonométrie à connaître...), on aurait alors :* $\lim\limits_{n\to+\infty} \cos n = \ell - \ell = 0$.

Ensuite, puisque $\cos(n+1) - \cos(n-1) = -2\sin 1 \sin n$, *on en déduit* $\lim\limits_{n\to+\infty} \sin n = 0$.

Mais on ne peut pas avoir simultanément $\lim\limits_{n\to+\infty} \sin n = \lim\limits_{n\to+\infty} \cos n = 0$, *puisque* $\sin^2 n + \cos^2 n = 1$, *d'où la contradiction cherchée.*

SF11.2 Majorer les sommes partielles

Dans certains cas, assez peu fréquents, il peut être plus facile de majorer directement les sommes partielles que le terme général de la série. Si l'on peut montrer que ces sommes partielles sont majorées (pour une série à termes positifs) cela prouvera la convergence de la série.

Exemple

Soit $\sum\limits_{n\geqslant 1} u_n$ une série à termes réels positifs, convergente. Montrer que la série $\sum\limits_{n\geqslant 1} \frac{\sqrt{u_n}}{n}$ converge.

Il faut ici penser à utiliser l'inégalité de Cauchy-Schwarz qui, dans $\mathbb{R}^n$ *muni de sa structure euclidienne canonique, s'écrit :*

$$\forall\,(x_1,\dots,x_n),\,(y_1,\dots,y_n)\in\mathbb{R}^n\ ,\ \left|\sum_{k=1}^{n} x_k y_k\right| \leqslant \sqrt{\sum_{k=1}^{n} x_k^2}\sqrt{\sum_{k=1}^{n} y_k^2}\,.$$

En l'appliquant à $x_k = \sqrt{u_k}$ *et* $y_k = \dfrac{1}{k}$ *on obtient :*

$$\sum_{k=1}^{n} \frac{\sqrt{u_k}}{k} \leqslant \sqrt{\sum_{k=1}^{n} u_k}\sqrt{\sum_{k=1}^{n} \frac{1}{k^2}} \leqslant \sqrt{\sum_{k=1}^{+\infty} u_k}\sqrt{\sum_{k=1}^{+\infty} \frac{1}{k^2}}\,,$$

la dernière inégalité étant justifiée par la convergence des deux séries écrites.

Les sommes partielles de la série de terme général positif $\dfrac{\sqrt{u_n}}{n}$ *étant majorées, cette série est convergente.*

SF11.3 Majorer ou minorer le terme général

Si $\sum u_n$ est une série à termes positifs, on peut essayer de majorer le terme général u_n par celui d'une série convergente afin de prouver sa convergence, ou le minorer par celui d'une série à termes positifs divergente afin de prouver sa divergence.

Exemple

Étudier la nature de la série de terme général $u_n = \dfrac{1}{(n!)^{1/n}}$ $(n \geqslant 1)$.

Puisque $n! = n.(n-1)\ldots.3.2 \leqslant n^n$, *on a* $(n!)^{1/n} \leqslant n$ *donc* $u_n \geqslant \dfrac{1}{n}$.

La série $\sum_{n\in\mathbb{N}^*} \dfrac{1}{n}$ *étant une série de Riemann divergente, la série* $\sum_{n\in\mathbb{N}^*} u_n$ *diverge aussi.*

Exemple

Étudier la nature de la série de terme général $u_n = \dfrac{\operatorname{Arctan} n}{(1+\cos n+\sqrt{n})^n}$.

On remarque d'abord que u_n *est bien positif.*

De plus, pour tout $n \in \mathbb{N}$, $\operatorname{Arctan} n \leqslant \dfrac{\pi}{2}$ *et* $1+\cos n+\sqrt{n} \geqslant 2$.

Donc pour $n \in \mathbb{N}$, $0 \leqslant u_n \leqslant \dfrac{\pi}{2^{n+1}}$.

Par comparaison à une série géométrique de raison $\dfrac{1}{2}$, *cela prouve la convergence de la série* $\sum u_n$.

✓ Il n'est pas toujours facile de trouver une majoration ou une minoration efficace. C'est pourquoi, dans de nombreux cas, les règles de comparaison qui vont suivre sont plus faciles à mettre en oeuvre.
Cependant, on pensera à utiliser cette méthode lorsque le terme général a vraiment une allure peu engageante, comme dans l'exemple précédent.

SF11.4 Utiliser un équivalent

L'utilisation des équivalents est fondamentale lors de l'étude d'une série ; en effet, même lorsqu'elle ne permet pas de conclure directement, elle permet au moins de simplifier l'expression du terme général avant d'appliquer éventuellement une autre méthode. Il s'agit de LA méthode à ne pas oublier (s'il n'y en avait qu'une !).

Exemple

Étudier la nature de la série de terme général $u_n = \dfrac{1}{n^\alpha}\left(\sqrt[3]{n+1}-\sqrt[3]{n}\right)$ selon les valeurs de α.

On utilise LA technique de base pour se ramener aux développements limités et équivalents usuels, à savoir factoriser par le terme prépondérant, en écrivant donc $n+1 = n\left(1+\dfrac{1}{n}\right)$:

$$u_n = \frac{1}{n^\alpha}\left(\sqrt[3]{n}\sqrt[3]{1+\frac{1}{n}}-\sqrt[3]{n}\right) = \frac{1}{n^{\alpha-1/3}}\left(\left(1+\frac{1}{n}\right)^{1/3}-1\right)$$

donc $u_n \underset{n\to+\infty}{\sim} \dfrac{1}{n^{\alpha-1/3}}\cdot\dfrac{1}{3n} = \dfrac{1}{3n^{\alpha+2/3}}$.

Par comparaison à une série de Riemann, la série $\sum u_n$ *converge si et seulement si* $\alpha+\dfrac{2}{3} > 1$ *soit* $\alpha > \dfrac{1}{3}$.

✓ Il faut bien sûr connaître parfaitement les équivalents et développements limités des fonctions usuelles, ainsi que les techniques de base pour les obtenir. Nous ne saurions trop vous recommander de revoir le cours de 1re année à ce sujet.

Parmi les équivalents à connaître, il ne faut pas négliger la formule de Stirling :

$$n! \underset{n\to+\infty}{\sim} \left(\frac{n}{\mathrm{e}}\right)^n \sqrt{2\pi n}\,.$$

Exemple

Étudier la nature de la série de terme général $u_n = \left(\dfrac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\right)^{\alpha}$.

On commence par simplifier l'expression à l'aide de factorielles :

$$\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n} = \frac{(2n)!}{\left(2\cdot 4\cdots 2n\right)^2} = \frac{(2n)!}{\left(2^n n!\right)^2}\,,$$

puis la formule de Stirling donne :

$$\frac{(2n)!}{\left(2^n n!\right)^2} \underset{n\to+\infty}{\sim} \frac{\left(\frac{2n}{\mathrm{e}}\right)^{2n}\sqrt{4\pi n}}{2^{2n}\left(\frac{n}{\mathrm{e}}\right)^{2n}(2\pi n)} = \frac{1}{\sqrt{\pi n}}\cdot$$

On en déduit $u_n \underset{n\to+\infty}{\sim} \dfrac{1}{(\pi n)^{\alpha/2}}$, *et par comparaison à une série de Riemann, la série* $\sum u_n$ *converge si et seulement si* $\alpha > 2$.

SF11.5 Utiliser des o et des O

On utilise principalement les o et les O pour comparer une série $\sum u_n$ à termes positifs (ou à termes complexes quelconques, en considérant $|u_n|$), à une série de Riemann $\sum \dfrac{1}{n^\alpha}$. Pour cela, on peut être amené à étudier la limite $\lim\limits_{n\to+\infty} n^\alpha u_n$ en discutant selon les valeurs de α.

✓ Certains d'entre vous connaissent peut-être la « règle de Riemann », ou « règle $n^\alpha u_n$ » ; cependant cette règle ne figure pas au programme officiel et il faut donc détailler le raisonnement à chaque fois.

Cette technique est à utiliser lorsque la recherche d'un équivalent simple n'est pas possible, par exemple si l'expression de u_n contient des logarithmes ou des exponentielles, qui se comparent facilement à n^α en $+\infty$.

Exemple

Étudier la nature de la série de terme général $P(n)\mathrm{e}^{-\beta n}$, où P est un polynôme non nul, en discutant selon la valeur du réel β.

On commence déjà par simplifier le problème en remarquant que, si le terme dominant de $P(X)$ *est* $a_p X^p$, *on a :* $|u_n| \underset{n\to+\infty}{\sim} |a_p|\, n^p \mathrm{e}^{-\beta n}$.

On en déduit, d'après les croissances comparées des suites usuelles :

- *si* $\beta \leqslant 0$, $|u_n|$ *ne tend pas vers* 0 *quand* $n \to +\infty$, *donc la série diverge grossièrement ;*

– *si* $\beta > 0$, $\lim\limits_{n\to+\infty} n^2 |u_n| = \lim\limits_{n\to+\infty} |a_p|\, n^{p+2} \mathrm{e}^{-\beta n} = 0$, *donc* $u_n = o\left(\dfrac{1}{n^2}\right)$ *et la série* $\sum u_n$ *est absolument convergente.*

Exemple

Étudier la nature de la série de terme général $u_n = \dfrac{1}{\operatorname{sh}\sqrt{\ln n}}$.

On commence déjà par simplifier le problème : puisque $\operatorname{sh} x = \dfrac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}$, *on a* $\operatorname{sh} x \underset{x\to+\infty}{\sim} \dfrac{\mathrm{e}^x}{2}$ *donc* $u_n \underset{n\to+\infty}{\sim} 2\mathrm{e}^{-\sqrt{\ln n}}$.
Soit $\alpha \in \mathbb{R}$. *On a* $n^\alpha u_n = \mathrm{e}^{\alpha \ln n} u_n \underset{n\to+\infty}{\sim} 2\mathrm{e}^{\alpha \ln n - \sqrt{\ln n}}$.
On constate alors que, pour $\alpha = 1$ *(par exemple),* $\lim\limits_{n\to+\infty} nu_n = +\infty$. *On aura donc* $nu_n \geqslant 1$ *pour* n *assez grand, soit* $u_n \geqslant \dfrac{1}{n}$, *et par comparaison à la série harmonique, la série* $\sum u_n$ *diverge.*

SF11.6 Utiliser la méthode de comparaison avec une intégrale

L'exemple suivant utilise directement le théorème **[S11.7]**.

Exemple

Étudier la série de terme général $u_n = \dfrac{1}{n(\ln n)^\alpha}$ $(n \geqslant 2)$ selon les valeurs de α.

Il est important de déjà remarquer que, si $\alpha \leqslant 0$, *alors* $u_n \geqslant \dfrac{1}{n}$, *donc la série* $\sum\limits_{n\geqslant 2} u_n$ *est alors divergente.*
On peut donc se limiter au cas $\alpha > 0$. *Ici le terme général s'écrit* $u_n = f(n)$ *avec* $f : t \mapsto \dfrac{1}{t(\ln t)^\alpha}$. f *est positive et continue sur* $]1\,;+\infty[$ *et, puisque* $\alpha > 0$, *elle y est décroissante (nul besoin de calcul de dérivée !).*
Le théorème de comparaison série-intégrale permet d'affirmer directement que la série $\sum\limits_{n\geqslant 2} u_n$ *converge si et seulement si l'intégrale* $\int_2^{+\infty} \dfrac{\mathrm{d}t}{t(\ln t)^\alpha}$ *existe.*
Or :
– *si* $\alpha \neq 1$, $\displaystyle\int_2^X \frac{\mathrm{d}t}{t(\ln t)^\alpha} = \left[\frac{1}{(1-\alpha)(\ln t)^{\alpha-1}}\right]_{t=2}^X$;
– *si* $\alpha = 1$, $\displaystyle\int_2^X \frac{\mathrm{d}t}{t\ln t} = \big[\ln(\ln t)\big]_2^X$.
Donc $\lim\limits_{X\to+\infty} \displaystyle\int_2^X f(t)\,\mathrm{d}t$ *existe si et seulement si* $\alpha > 1$.
En conclusion, la série $\sum\limits_{n\geqslant 2} \dfrac{1}{n(\ln n)^\alpha}$ *converge si et seulement si* $\alpha > 1$.

✓ Ce résultat, qui est un cas particulier des *séries de Bertrand*, est souvent utilisé ; mais ne faisant pas partie du programme officiel, vous devez savoir le redémontrer rapidement.

L'exemple suivant montre comment utiliser la méthode de comparaison avec une intégrale, même dans le cas d'une fonction croissante. Il faut savoir refaire ce genre

de calculs, dont le principe est simple mais qui demande un peu de technique et du soin, en particulier dans la manipulation des indices.

Exemple

Étudier la série de terme général $u_n = \dfrac{\ln(n!)}{n^\alpha}$ $(n \geqslant 1)$ selon les valeurs de α.

Pour $n \in \mathbb{N}^*$, posons $v_n = \ln(n!) = \displaystyle\sum_{k=2}^{n} \ln k$. La fonction $\ln$ étant croissante sur $[1\,;+\infty[$, on a :

$$\forall\, k \geqslant 2\ ,\ \int_{k-1}^{k} \ln t\,\mathrm{d}t \leqslant \ln k \leqslant \int_{k}^{k+1} \ln t\,\mathrm{d}t\,.$$

En sommant ces inégalités pour k variant de 2 à n, on obtient :

$$\forall\, n \geqslant 2\ ,\ \int_{1}^{n} \ln t\,\mathrm{d}t \leqslant v_n \leqslant \int_{2}^{n+1} \ln t\,\mathrm{d}t\,.$$

Une primitive de $\ln$ étant $t \mapsto t\ln t - t$, on en tire :

$$\forall\, n \geqslant 2\ ,\ n\ln n - n + 1 \leqslant v_n \leqslant (n+1)\ln(n+1) - (n+1) - 2\ln 2 + 2\,,$$

d'où par encadrement : $\displaystyle\lim_{n\to+\infty} \frac{v_n}{n\ln n} = 1$, soit : $v_n \underset{n\to+\infty}{\sim} n\ln n$.

Ainsi, $u_n \underset{n\to+\infty}{\sim} \dfrac{\ln n}{n^{\alpha-1}}$. On distingue alors deux cas.

- *Si $\alpha - 1 \leqslant 1$, $\dfrac{\ln n}{n^{\alpha-1}} \geqslant \dfrac{\ln n}{n} \geqslant \dfrac{1}{n}$ pour $n \geqslant 3$. Par comparaison à la série harmonique, la série $\sum \dfrac{\ln n}{n^{\alpha-1}}$ diverge, et puisqu'il s'agit d'une série à termes positifs, $\sum u_n$ aussi.*
- *Si $\alpha - 1 > 1$, soit β tel que $\alpha - 1 > \beta > 1$. Alors $\displaystyle\lim_{n\to+\infty} n^\beta \frac{\ln n}{n^{\alpha-1}} = 0$, donc $\dfrac{\ln n}{n^{\alpha-1}} = o\Big(\dfrac{1}{n^\beta}\Big)$, et puisque $\beta > 1$, la série $\sum \dfrac{\ln n}{n^{\alpha-1}}$ converge, donc $\sum u_n$ aussi.*

Conclusion : la série $\sum u_n$ converge si et seulement si $\alpha - 1 > 1$ soit $\alpha > 2$.

✓ Là encore, nous avons utilisé à la fin de l'exemple la méthode classique d'étude d'une série de Bertrand.

Le dernier exemple que nous présentons est moins classique, et peut s'appliquer à des fonctions plus variées. La méthode consiste à utiliser la relation suivante, obtenue pour une fonction f de classe $\mathscr{C}^1$ sur un intervalle de la forme $[a\,;+\infty[$, par intégration par parties :

$$\int_{n}^{n+1} f(t)\,\mathrm{d}t = \underbrace{\Big[\big(t-(n+1)\big)f(t)\Big]_{n}^{n+1}}_{=f(n)} - \int_{n}^{n+1} \big(t-(n+1)\big)f'(t)\,\mathrm{d}t\,.$$

✓ L'astuce consiste ici, dans l'intégration par parties, à utiliser comme primitive de 1 la fonction $t \mapsto t-(n+1)$.

Cette égalité permet alors de relier la convergence de la série $\sum f(n)$ à celle d'intégrales.

Exemple

Montrer que la série de terme général $u_n = \dfrac{\sin(\ln n)}{n}$ $(n \geqslant 1)$ est divergente.

Pour $t > 0$, on pose $f(t) = \dfrac{\sin(\ln t)}{t}$. f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$ et le calcul ci-dessus donne :

$$\int_n^{n+1} \frac{\sin(\ln t)}{t}\,\mathrm{d}t = \frac{\sin(\ln n)}{n} - \int_n^{n+1} \big(t-(n+1)\big)\frac{\cos(\ln t)-\sin(\ln t)}{t^2}\,\mathrm{d}t\,.$$

Posons $v_n = \displaystyle\int_n^{n+1} \big(t-(n+1)\big)\frac{\cos(\ln t)-\sin(\ln t)}{t^2}\,\mathrm{d}t$.

La série de terme général v_n est absolument convergente puisque :

$$|v_n| \leqslant \int_n^{n+1} \underbrace{|t-(n+1)|}_{\leqslant 1} \frac{|\cos(\ln t)-\sin(\ln t)|}{t^2}\,\mathrm{d}t \leqslant 2\int_n^{n+1}\frac{\mathrm{d}t}{t^2} = \frac{2}{n(n+1)}\,.$$

Il en résulte que la série de terme général $\dfrac{\sin(\ln n)}{n}$ est de même nature que celle de terme général $\displaystyle\int_n^{n+1}\frac{\sin(\ln t)}{t}\,\mathrm{d}t$. Mais, par le changement de variables $u = \ln t$:

$$\int_n^{n+1} \frac{\sin(\ln t)}{t}\,\mathrm{d}t = \int_{\ln n}^{\ln(n+1)} \sin u\,\mathrm{d}u = -\cos(\ln(n+1)) + \cos(\ln n)\,,$$

et cette série télescopique diverge (car la suite $n \mapsto \cos(\ln n)$ diverge).

Il en est donc de même de la série $\sum u_n$.

✓ On notera ici l'intérêt de l'intégration par parties : elle permet d'augmenter le degré du dénominateur, donc d'introduire une série convergente. Cette technique est importante à retenir, nous en verrons d'autres exemples plus loin, par exemple en **[SF14.4]**.

SF11.7 Utiliser la règle de d'Alembert

Cette règle n'est pas très précise : elle ne permet pas, par exemple, d'étudier la convergence des séries de Riemann. Son utilisation doit être limitée à l'étude de séries dont le terme général s'exprime sous forme d'un produit ou d'un quotient, ou contient des factorielles ou des puissances.

Exemple

Étudier la nature de la série de terme général $u_n = \dfrac{(n!)^3}{n^n(2n)!}$.

Il s'agit bien d'une série à termes strictement positifs. On calcule :

$$\frac{u_{n+1}}{u_n} = \frac{n^n}{(n+1)^{n+1}}\cdot\frac{(n+1)!^3}{n!^3}\cdot\frac{(2n)!}{(2n+2)!}$$
$$= \frac{n^n(n+1)^3}{(n+1)^{n+1}(2n+1)(2n+2)} = \left(\frac{n}{n+1}\right)^n \frac{n+1}{2(2n+1)}\,.$$

Or $\displaystyle\lim_{n\to+\infty}\left(\frac{n+1}{n}\right)^n = \lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to+\infty}\mathrm{e}^{n\ln\left(1+\frac{1}{n}\right)} = \mathrm{e}$ *donc* $\displaystyle\lim_{n\to+\infty}\frac{u_{n+1}}{u_n} = \frac{1}{4\mathrm{e}}$. *La limite obtenue étant < 1, la série converge.*

SF11.8 Utiliser la convergence absolue

Dès que l'on doit étudier une série à termes réels pas toujours positifs, à termes complexes ou plus généralement à valeurs dans un espace vectoriel normé, on peut penser à étudier la convergence absolue. Cela permet d'utiliser les critères spécifiques de convergence pour les séries à termes positifs. Et même lorsque la série est alternée, l'étude de la convergence absolue peut être parfois plus rapide que la vérification des hypothèses du critère spécial.

Exemple

Étudier la nature de la série de terme général $u_n = \left(1 - \frac{n}{\ln n}\right)^{-n}$.

On remarque déjà que, pour $n \geqslant 2$, $n \geqslant \ln n$ donc u_n est du signe de $(-1)^n$. On a donc $|u_n| = \left(\frac{n}{\ln n} - 1\right)^{-n}$. Or $\lim\limits_{n\to+\infty} \left(\frac{n}{\ln n} - 1\right) = +\infty$ (croissances comparées) donc pour n assez grand, on a $\frac{n}{\ln n} - 1 \geqslant 2$ d'où $|u_n| \leqslant \frac{1}{2^n}$.
Par comparaison à une série géométrique de raison $\frac{1}{2}$, la série $\sum u_n$ est absolument convergente, donc convergente.

SF11.9 Utiliser directement le critère spécial des séries alternées

Le critère spécial sur les séries alternées est très simple à utiliser, mais il faut en vérifier soigneusement les hypothèses. Beaucoup trop d'entre vous pensent qu'il s'agit d'une série alternée dès qu'il y a un $(-1)^n$ en facteur, ou considèrent comme évident la décroissance de la valeur absolue du terme général.
Dans beaucoup de cas, une vérification sérieuse s'impose.

Exemple

Étudier, en discutant selon les paramètres réels a et b, la série de terme général :

$$u_n = (-1)^n \frac{n^a}{(n+1)^b}.$$

- *Lorsque $n \to +\infty$ on a $|u_n| \sim n^{a-b}$ donc, si $a \geqslant b$, le terme général ne tend pas vers 0 et la série diverge grossièrement.*
- *Supposons désormais $a < b$.*
  - *u_n est du signe de $(-1)^n$ donc la suite u est bien alternée.*
  - *L'équivalent trouvé précédemment montre que $\lim\limits_{n\to+\infty} u_n = 0$.*
  - *Pour pouvoir appliquer le critère, il nous reste à montrer que la suite $n \mapsto \frac{n^a}{(n+1)^b}$ est décroissante. Pour cela, on considère la fonction $f : x \in \mathbb{R}_+ \mapsto \frac{x^a}{(x+1)^b}$. f est dérivable sur $\mathbb{R}_+$ et :*

$$\forall\, x \geqslant 0,\ f'(x) = \frac{x^{a-1}\left((a-b)x + a\right)}{(x+1)^{b+1}}.$$

$f'(x)$ est donc du signe de $(a-b)x + a$ et, puisque $a - b$ est strictement négatif, $f'(x)$ est négative pour $x > \frac{a}{b-a}$.

Puisque $|u_n| = f(n)$, il en résulte que la suite $(|u_n|)$ est décroissante à partir d'un certain rang.

Le théorème des séries alternées s'applique donc (à partir d'un certain rang), ce qui montre que dans le cas $a < b$ la série converge.

SF11.10 Utiliser des développements limités

Dans les cas où un simple équivalent ne peut suffire pour conclure, ou dans le cas de séries alternées où le critère spécial ne s'applique pas, on peut effectuer un développement limité du terme général afin d'écrire la série comme somme de séries connues.

Exemple

Nature de la série de terme général $u_n = \dfrac{(-1)^n}{\sqrt{n}+(-1)^n}$ $(n \geqslant 2)$.

On effectue un développement limité de u_n à un ordre suffisant pour trouver le terme général d'une série absolument convergente (ici, un $O\left(\frac{1}{n^{3/2}}\right)$) :

$$u_n = \frac{(-1)^n}{\sqrt{n}} \frac{1}{1+\frac{(-1)^n}{\sqrt{n}}} = \frac{(-1)^n}{\sqrt{n}} \left(1 - \frac{(-1)^n}{\sqrt{n}} + O\left(\frac{1}{n}\right)\right)$$
$$= \frac{(-1)^n}{\sqrt{n}} - \frac{1}{n} + O\left(\frac{1}{n^{3/2}}\right).$$

La série de terme général u_n est donc la somme de 3 séries :

- *celle de terme général $\dfrac{(-1)^n}{\sqrt{n}}$, qui est convergente d'après le critère spécial sur les séries alternées ;*
- *celle de terme général $O\left(\frac{1}{n^{3/2}}\right)$, qui est absolument convergente par comparaison à une série de Riemann ;*
- *et enfin celle de terme général $\dfrac{1}{n}$, qui est divergente.*

Il en résulte que $\sum u_n$ est divergente.

✓ Dans cet exemple, on avait $u_n \underset{n\to+\infty}{\sim} \dfrac{(-1)^n}{\sqrt{n}}$; cependant, la série $\sum u_n$ diverge alors que la série $\sum \dfrac{(-1)^n}{\sqrt{n}}$ converge. Cela montre que la règle de comparaison par équivalence ne peut s'appliquer à des séries dont le terme général n'est pas de signe constant !

Exemple

Déterminer les réels a et b de sorte que la série de terme général :

$$u_n = \sqrt{n} + a\sqrt{n+1} + b\sqrt{n+2}$$

soit convergente. Calculer alors sa somme.

On effectue un développement limité de u_n, à un ordre convenable :

$$\begin{aligned} u_n &= \sqrt{n}\left(1 + a\left(1+\frac{1}{n}\right)^{1/2} + b\left(1+\frac{2}{n}\right)^{1/2}\right) \\ &= \sqrt{n}\left(1 + a + b + \frac{a+2b}{2n} + O\left(\frac{1}{n^2}\right)\right) \\ &= (1+a+b)\sqrt{n} + \frac{a+2b}{\sqrt{n}} + O\left(\frac{1}{n^{3/2}}\right). \end{aligned}$$

On en déduit facilement, comme dans l'exemple précédent, que la série de terme général u_n converge si et seulement si $1+a+b = a+2b = 0$, ce qui équivaut à $a = -2$ et $b = 1$.

Dans ce cas, $u_n = \sqrt{n} - 2\sqrt{n+1} + \sqrt{n+2}$ et, après télescopage :

$$\forall N \in \mathbb{N}\ ,\ \sum_{n=0}^{N} u_n = -1 + \sqrt{N+2} - \sqrt{N+1}$$

et puisque $\lim\limits_{N\to+\infty}\left(\sqrt{N+2} - \sqrt{N+1}\right) = 0$ (multiplier par la quantité conjuguée par exemple), on en déduit $\sum\limits_{n=0}^{+\infty} u_n = -1$.

SF11.11 Utiliser des séries télescopiques

Outre le calcul de la somme de certaines séries, les séries télescopiques peuvent aussi servir à étudier certaines suites, comme il a été rappelé en **[S11.2]** et **[SF8.4]**.

Exemple

Pour $n \in \mathbb{N}$ on pose $u_n = \dfrac{1}{n!}\sqrt{n}\left(\dfrac{n}{\mathrm{e}}\right)^n$ et $v_n = \ln u_{n+1} - \ln u_n$.

En étudiant la série $\sum v_n$, montrer qu'il existe une constante $C > 0$ telle que $n! \underset{n\to+\infty}{\sim} C\sqrt{n}\left(\dfrac{n}{\mathrm{e}}\right)^n$.

On calcule déjà : $\dfrac{u_{n+1}}{u_n} = \dfrac{1}{n+1}\sqrt{\dfrac{n+1}{n}}\left(\dfrac{n+1}{\mathrm{e}}\right)^{n+1}\left(\dfrac{\mathrm{e}}{n}\right)^n = \dfrac{1}{\mathrm{e}}\left(\dfrac{n+1}{n}\right)^{n+1/2}$.

Un développement limité donne alors :

$$\begin{aligned} v_n &= \ln\left(\frac{u_{n+1}}{u_n}\right) = -1 + \left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right) \\ &= -1 + \left(n+\frac{1}{2}\right)\left(\frac{1}{n} - \frac{1}{2n^2} + O\left(\frac{1}{n^3}\right)\right) = O\left(\frac{1}{n^2}\right), \end{aligned}$$

ce qui implique par comparaison que la série $\sum v_n$ converge absolument, donc converge, donc la suite $(\ln u_n)$ converge aussi.

En notant ℓ sa limite, on a donc $\lim\limits_{n\to+\infty} \ln u_n = \ell$ soit $\lim\limits_{n\to+\infty} u_n = \mathrm{e}^{\ell}$.

En posant $\mathrm{e}^{\ell} = \dfrac{1}{C}$, avec $C > 0$, on obtient l'équivalent cherché.

✓ On peut démontrer que $C = \sqrt{2\pi}$, ce qui fournit la *formule de Stirling* :

$$n! \underset{n\to+\infty}{\sim} \left(\frac{n}{\mathrm{e}}\right)^n \sqrt{2\pi n}.$$

SF11.12 Revenir à la définition avec les sommes partielles

Exemple

Soit σ une permutation de $\mathbb{N}^*$. Pour $n \in \mathbb{N}^*$, soit $S_n = \sum_{k=1}^{n} \frac{\sigma(k)}{k^2}$.

Montrer que : $\forall\, n \in \mathbb{N}^*$, $S_{2n} - S_n \geqslant \frac{1}{8}$. En déduire la nature de la série $\sum \frac{\sigma(n)}{n^2}$.

- $S_{2n} - S_n = \sum_{k=n+1}^{2n} \sigma(k) \underbrace{\left(\frac{1}{k}\right)^2}_{\geqslant \frac{1}{2n}} \geqslant \frac{1}{4n^2} \sum_{k=n+1}^{2n} \sigma(k)$.

  Puisque σ est injective, les n entiers $\sigma(n+1), \ldots, \sigma(2n)$ ont une somme plus grande que $1 + 2 + \ldots + n = \frac{n(n+1)}{2}$ d'où :

  $$S_{2n} - S_n \geqslant \frac{n(n+1)}{8n^2} = \frac{n+1}{8n} \geqslant \frac{1}{8}.$$

- *Supposons maintenant que la suite (S_n) converge. On peut alors passer à la limite dans l'inégalité précédente et l'on obtient : $0 \geqslant \frac{1}{8}$, ce qui est évidemment absurde. Bref, la suite (S_n) diverge, c'est-à-dire que la série $\sum \frac{\sigma(n)}{n^2}$ diverge.*

Il est aussi indispensable de revenir à la définition lorsque l'on veut faire des *sommations par paquets*, comme dans l'exemple suivant.

Exemple

Nature de la série de terme général $x^{p(n)}$ où $p(n)$ est le nombre de chiffres dans l'écriture de n en base 10.

On remarque déjà que le terme général de cette série ne tend pas vers 0 lorsque $|x| \geqslant 1$, donc dans ce cas la série diverge grossièrement.

On suppose donc pour la suite $|x| < 1$. L'idée est de regrouper les termes $u_n = x^{p(n)}$ pour lesquels $p(n)$ possède la même valeur.
Or dire que le nombre de chiffres de l'écriture de n en base 10 est égal à $k+1$ (avec $k \in \mathbb{N}$) équivaut à l'encadrement : $10^k \leqslant n < 10^{k+1}$. Pour simplifier les écritures, notons $n_k = 10^k - 1$.
En notant $S_n = \sum_{k=0}^{n} u_k$, on aura donc, pour tout entier k :

$$S_{n_{k+1}} - S_{n_k} = \sum_{n=n_k+1}^{n_{k+1}} u_n = \sum_{n=n_k+1}^{n_{k+1}} x^{k+1} = (n_{k+1} - n_k)\, x^{k+1}$$
$$= (10^{k+1} - 10^k)\, x^{k+1} = 9.10^k x^{k+1} = \frac{9}{10}(10x)^{k+1}.$$

Or la série de terme général $(10x)^{k+1}$, qui est une série géométrique, converge si et seulement si $|x| < \frac{1}{10}$; il en est donc de même de la série télescopique $\sum_{k \in \mathbb{N}} (S_{n_{k+1}} - S_{n_k})$, donc de la suite $(S_{n_k})_{k \in \mathbb{N}}$.

Pour tout $n \in \mathbb{N}$, il existe un unique $k \in \mathbb{N}$ tel que $10^k \leqslant n < 10^{k+1}$ et l'on a alors $S_{n_k} \leqslant S_n \leqslant S_{n_{k+1}}$; donc la suite $(S_n)_{n\in\mathbb{N}}$ est majorée si et seulement si la suite $(S_{n_k})_{k\in\mathbb{N}}$ l'est.
Puisqu'il s'agit d'une série à termes positifs, on en conclut que le série $\sum x^{p(n)}$ converge si et seulement si $|x| < \dfrac{1}{10}$.

Un autre cas où l'utilisation des sommes partielles est incontournable est l'utilisation de la *transformation d'Abel*.
On ne peut pas vous demander d'appliquer cette méthode directement puisqu'elle ne figure plus dans le programme officiel, mais elle est très performante et il faut essayer de bien comprendre l'exemple qui suit.

Exemple

Démontrer que, pour tout $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, la série $\displaystyle\sum_{n\geqslant 1} \frac{\mathrm{e}^{\mathrm{i}n\theta}}{n}$ est convergente.

Notons $S_n = \displaystyle\sum_{k=1}^{n} \frac{\mathrm{e}^{\mathrm{i}k\theta}}{k}$ les sommes partielles de la série. L'idée de la transformation d'Abel est d'introduire les sommes : $B_n = \displaystyle\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}k\theta}$, puis d'exprimer les S_n en fonction des B_k en utilisant la relation $\mathrm{e}^{\mathrm{i}k\theta} = B_k - B_{k-1}$, valable pour tout entier $k \geqslant 1$ sous réserve de poser $B_0 = 0$.
On a ainsi, pour tout $n \geqslant 1$:

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \frac{\mathrm{e}^{\mathrm{i}k\theta}}{k} = \sum_{k=1}^{n} \frac{B_k - B_{k-1}}{k} = \sum_{k=1}^{n} \frac{B_k}{k} - \sum_{k=1}^{n} \frac{B_{k-1}}{k} \\ &= \sum_{k=1}^{n} \frac{B_k}{k} - \sum_{k=0}^{n-1} \frac{B_k}{k+1} \quad \textit{(changement d'indice dans la 2ème somme)} \\ &= \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) B_k + \frac{B_n}{n+1} \quad \textit{(car } B_0 = 0\textit{)} \quad (*) \end{aligned}$$

Or B_n est la somme des n premiers termes d'une suite géométrique de raison $\mathrm{e}^{\mathrm{i}\theta} \neq 1$ (d'après l'hypothèse $\theta \notin 2\pi\mathbb{Z}$), et de premier terme $\mathrm{e}^{\mathrm{i}\theta}$, donc :

$$B_n = \mathrm{e}^{\mathrm{i}\theta} \frac{1 - \mathrm{e}^{\mathrm{i}n\theta}}{1 - \mathrm{e}^{\mathrm{i}\theta}},$$

d'où : $|B_n| = \underbrace{\left|\mathrm{e}^{\mathrm{i}\theta}\right|}_{=1} \dfrac{\left|1 - \mathrm{e}^{\mathrm{i}n\theta}\right|}{|1 - \mathrm{e}^{\mathrm{i}\theta}|} \leqslant \dfrac{1 + \left|\mathrm{e}^{\mathrm{i}n\theta}\right|}{|1 - \mathrm{e}^{\mathrm{i}\theta}|} = \dfrac{2}{|1 - \mathrm{e}^{\mathrm{i}\theta}|}.$

La suite (B_n) est donc bornée. Il en résulte que $\displaystyle\lim_{n\to+\infty} \frac{B_n}{n+1} = 0$ et d'après le résultat de la transformation d'Abel $(*)$, la série de terme général $\dfrac{\mathrm{e}^{\mathrm{i}k\theta}}{k}$ est de même nature que celle de terme général $\left(\dfrac{1}{k} - \dfrac{1}{k+1}\right) B_k$.
Pour conclure, il suffit de remarquer que cette série est absolument convergente puisque $\left|B_k \left(\dfrac{1}{k} - \dfrac{1}{k+1}\right)\right| = \dfrac{|B_k|}{k(k+1)} = O\left(\dfrac{1}{k^2}\right)$ (en utilisant le fait que la suite (B_k) est bornée), et d'après la règle de comparaison **[S11.6]**.

■ Comment calculer la somme d'une série ?

Le nombre de séries dont on sait effectivement calculer la somme est très limité, donc les méthodes que l'on peut utiliser sont assez peu nombreuses. Nous en présentons certaines ci-dessous.
Il existe également des méthodes utilisant les séries de fonctions (dérivation et intégration, séries entières), que nous verrons plus tard.

SF11.13 Utiliser des séries connues

Il s'agit ici de mettre le terme général étudié sous la forme d'une combinaison linéaire de termes généraux de séries de sommes connues. Cela exige bien sûr de *connaître parfaitement toutes les formules de sommes de séries classiques.*

Exemple

Calculer $S = \sum_{n=0}^{+\infty} \frac{n^2+5n-2}{n!}$.

On commence par démontrer, à l'aide de la règle de d'Alembert (nous ne reproduisons pas les calculs ici), que toute série de la forme $\sum \frac{P(n)}{n!}$, où P est un polynôme, est convergente. Cela justifie les calculs ci-dessous :

$$\begin{aligned} S &= \sum_{n=0}^{+\infty} \frac{n^2+5n-2}{n!} = \sum_{n=0}^{+\infty} \frac{n(n-1)+6n-2}{n!} \\ &= \sum_{n=0}^{+\infty} \frac{n(n-1)}{n!} + 6\sum_{n=0}^{+\infty} \frac{n}{n!} - 2\sum_{n=0}^{+\infty} \frac{1}{n!} \\ &= \sum_{n=2}^{+\infty} \frac{1}{(n-2)!} + 6\sum_{n=1}^{+\infty} \frac{1}{(n-1)!} - 2\sum_{n=0}^{+\infty} \frac{1}{n!} = \mathrm{e} + 6\mathrm{e} - 2\mathrm{e} = 5\mathrm{e}. \end{aligned}$$

✓ *On notera l'importance de la transformation initiale qui a permis de simplifier les factorielles.*
Elle consiste, dans le calcul de la somme d'une série de la forme $\sum \frac{P(n)}{n!}$, où P est un polynôme de degré p, à exprimer le polynôme P dans la base $\big(1, X, X(X-1), \ldots, X(X-1)\ldots(X-p+1)\big)$ de $\mathbb{R}_p[X]$.

Exemple

Pour $|x| < 1$, calculer $\sum_{n=0}^{+\infty} x^n \cos n\theta$, où θ est un réel donné.

On a $|x^n \cos n\theta| \leqslant |x|^n$; la série géométrique $\sum |x|^n$, à termes positifs, étant convergente, cela prouve la convergence absolue de la série donnée. On peut donc écrire directement, puisque toutes les séries écrites convergent :

$$\sum_{n=0}^{+\infty} x^n \cos n\theta = \mathcal{R}e\left(\sum_{n=0}^{+\infty} x^n \mathrm{e}^{\mathrm{i}n\theta}\right) = \mathcal{R}e\left(\sum_{n=0}^{+\infty} \left(x\mathrm{e}^{\mathrm{i}\theta}\right)^n\right) = \mathcal{R}e\left(\frac{1}{1-x\mathrm{e}^{\mathrm{i}\theta}}\right)$$

donc $\displaystyle\sum_{n=0}^{+\infty} x^n \cos n\theta = \mathcal{R}e\left(\frac{1}{1-x\mathrm{e}^{\mathrm{i}\theta}}\right) = \mathcal{R}e\left(\frac{1-x\mathrm{e}^{-\mathrm{i}\theta}}{\left|1-x\mathrm{e}^{\mathrm{i}\theta}\right|^2}\right) = \frac{1-x\cos\theta}{x^2-2x\cos\theta+1}\cdot$

SF11.14 Revenir à la définition avec les sommes partielles

On introduit ici la suite des sommes partielles que l'on cherche à simplifier pour pouvoir facilement montrer qu'elle converge et déterminer sa limite.

Exemple

Pour $n \geqslant 2$, on pose : $u_n = \ln\left(1-\frac{1}{n^2}\right)$. Montrer que la série $\sum u_n$ converge et calculer sa somme.

Soit $n \geqslant 2$. On pose : $S_n = \displaystyle\sum_{k=2}^{n} u_k$. On a :

$$\begin{aligned} S_n &= \sum_{k=2}^{n} \ln\left(\frac{k^2-1}{k^2}\right) = \sum_{k=2}^{n} \ln\left(\frac{(k-1)(k+1)}{k^2}\right) \\ &= \sum_{k=2}^{n} \big(\ln(k-1) + \ln(k+1) - 2\ln k\big) \\ &= \sum_{k=2}^{n} \big(\ln(k-1) - \ln k\big) + \sum_{k=2}^{n} \big(\ln(k+1) - \ln k\big). \end{aligned}$$

On reconnaît deux sommes télescopiques, d'où :

$$S_n = \big(\ln 1 - \ln n\big) + \big(\ln(n+1) - \ln 2\big) = \ln\left(\frac{n+1}{n}\right) - \ln 2\,.$$

Donc $\lim\limits_{n\to+\infty} S_n = -\ln 2$ *: la série* $\sum u_n$ *converge et* $\displaystyle\sum_{n=2}^{+\infty} u_n = -\ln 2$.

✓ On aurait pu prouver directement la convergence de la série en utilisant $u_n \sim -\frac{1}{n^2}$, mais cela n'était guère utile pour la suite du calcul.

SF11.15 Utiliser le produit de Cauchy

Exemple

Pour $n \geqslant 2$, on pose : $u_n = \dfrac{\ln n}{2^n}$ et $v_n = \dfrac{1}{2^n}\ln\left(1-\dfrac{1}{n}\right)\cdot$

1) Montrer que les deux séries $\displaystyle\sum_{n\geqslant 2} u_n$ et $\displaystyle\sum_{n\geqslant 2} v_n$ sont convergentes.

2) En utilisant la relation $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{2^n} = 2$, montrer que : $\displaystyle\sum_{n=2}^{+\infty} u_n = -2\sum_{n=2}^{+\infty} v_n$.

1) *Puisque, d'après les croissances comparées des suites usuelles, on a :*

$$\lim_{n\to+\infty} n^2 u_n = \lim_{n\to+\infty} n^2 v_n = 0\,,$$

u_n et v_n sont tous deux des $o\left(\frac{1}{n^2}\right)$, ce qui prouve que les deux séries sont absolument convergentes par comparaison à une série de Riemann.

2) *La relation proposée par l'énoncé découle directement des formules sur la somme d'une série géométrique.*

Cette relation suggère de calculer le produit de Cauchy des deux séries $\sum_{n\geqslant 2} v_n$ *et* $\sum_{n\geqslant 0} \frac{1}{2^n}$. *Pour pouvoir utiliser directement les formules du cours, on posera* $v_0 = v_1 = 0$ *; ainsi,* $\sum_{n=2}^{+\infty} v_n = \sum_{n=0}^{+\infty} v_n$.

En notant w_n *le terme général de la série produit de Cauchy des deux séries* $\sum_{n\geqslant 0} v_n$ *et* $\sum_{n\geqslant 0} \frac{1}{2^n}$, *on a* $w_0 = w_1 = 0$ *et, pour* $n \geqslant 2$ *:*

$$w_n = \sum_{k=0}^{n} \frac{1}{2^{n-k}} v_k = \frac{1}{2^n} \sum_{k=2}^{n} \ln\left(1 - \frac{1}{k}\right) = \frac{1}{2^n} \sum_{k=2}^{n} \big(\ln(k-1) - \ln k\big) = -\frac{\ln n}{2^n}$$

par télescopage.

Les deux séries étant absolument convergentes, le théorème **[S11.10]** *permet d'écrire :*

$$\sum_{n=0}^{+\infty} w_n = \sum_{n=2}^{+\infty} -\frac{\ln n}{2^n} = \left(\sum_{n=0}^{+\infty} \frac{1}{2^n}\right)\left(\sum_{n=2}^{+\infty} v_n\right),$$

ce qui est le résultat demandé.

SF11.16 Utiliser une intégrale

Les théorèmes d'interversion série-intégrale que nous avons rappelés en **[S12.10]** et **[S14.10]** peuvent parfois permettre de calculer la somme d'une série, en se ramenant au calcul d'une intégrale. Nous en verrons l'utilisation un peu plus loin, mais donnons d'ores et déjà un exemple où l'on peut calculer la somme d'une série à l'aide d'une intégrale, sans avoir besoin d'utiliser ces théorèmes.

Exemple

En utilisant la relation $\int_0^1 t^n \,\mathrm{d}t = \frac{1}{n+1}$, prouver la convergence et calculer la somme de la série : $\sum_{n\geqslant 0} \frac{\mathrm{e}^{\mathrm{i}n\theta}}{n+1}$ pour $\theta \in]0\,;2\pi[$.

La démonstration directe de la convergence de la série a été faite dans un exemple page 464. Cependant, nous allons ici démontrer directement que la suite des sommes partielles converge, et calculer sa limite.
Avec l'indication de l'énoncé on a, pour tout $n \in \mathbb{N}$ *:*

$$S_n = \sum_{k=0}^{n} \frac{\mathrm{e}^{\mathrm{i}k\theta}}{k+1} = \int_0^1 \left(\sum_{k=0}^{n} \left(\mathrm{e}^{\mathrm{i}\theta} t\right)^k\right) \mathrm{d}t = \int_0^1 \frac{1 - \left(\mathrm{e}^{\mathrm{i}\theta} t\right)^{n+1}}{1 - t\mathrm{e}^{\mathrm{i}\theta}} \,\mathrm{d}t$$

$$= \int_0^1 \frac{1}{1 - t\mathrm{e}^{\mathrm{i}\theta}} \,\mathrm{d}t \; + R_n \quad \textit{(on a bien } t\mathrm{e}^{\mathrm{i}\theta} \neq 1 \textit{ pour tout } t \in [0\,;1]\textit{)}$$

avec $R_n = -\int_0^1 \frac{\left(\mathrm{e}^{\mathrm{i}\theta} t\right)^{n+1}}{1 - t\mathrm{e}^{\mathrm{i}\theta}} \,\mathrm{d}t$.

On a : $|R_n| \leqslant \displaystyle\int_0^1 \frac{t^{n+1}}{|1-te^{i\theta}|}\,dt$ *; or la fonction* $t \mapsto |1-te^{i\theta}|$ *ne s'annule pas sur le compact* $[0\,;1]$ *donc y admet un minimum* m *strictement positif donc :*

$$|R_n| \leqslant \frac{1}{m}\int_0^1 t^{n+1}\,dt = \frac{1}{m}\cdot\frac{1}{n+2} \underset{n\to+\infty}{\longrightarrow} 0\,,$$

ce qui prouve que la série converge et a pour somme $S(\theta) = \displaystyle\int_0^1 \frac{1}{1-te^{i\theta}}\,dt$. *Il ne reste plus qu'à calculer cette intégrale, ce qui est un peu long mais n'utilise que des techniques classiques, détaillées dans un prochain chapitre. On supposera dans un 1er temps* $\theta \in]0\,;\pi[$. *On calcule :*

$$\begin{aligned}
S(\theta) &= \int_0^1 \frac{1-te^{-i\theta}}{t^2-2t\cos\theta+1}\,dt = e^{-i\theta}\int_0^1 \frac{e^{i\theta}-t}{t^2-2t\cos\theta+1}\,dt\\
&= e^{-i\theta}\int_0^1 \frac{\cos\theta-t}{t^2-2t\cos\theta+1}\,dt + ie^{-i\theta}\int_0^1 \frac{\sin\theta}{(t-\cos\theta)^2+\sin^2\theta}\,dt\\
&= -\frac{e^{-i\theta}}{2}\left[\ln(t^2-2t\cos\theta+1)\right]_0^1 + ie^{-i\theta}\left[\operatorname{Arctan}\left(\frac{t-\cos\theta}{\sin\theta}\right)\right]_0^1\\
&= -\frac{e^{-i\theta}}{2}\ln\left(4\sin^2\frac{\theta}{2}\right) + ie^{-i\theta}\left(\operatorname{Arctan}\left(\tan\frac{\theta}{2}\right)+\operatorname{Arctan}(\cot\theta)\right).
\end{aligned}$$

Pour $0<\theta<\pi$, *on a* $\operatorname{Arctan}(\cot\theta) = \frac{\pi}{2}-\theta$, $\operatorname{Arctan}\left(\tan\frac{\theta}{2}\right) = \frac{\theta}{2}$ *et* $\sin\frac{\theta}{2} > 0$, *d'où :*

$$S(\theta) = e^{-i\theta}\left(-\ln\left(2\sin\frac{\theta}{2}\right) + i\left(\frac{\pi}{2}-\frac{\theta}{2}\right)\right) \quad (*)$$

Cette formule reste valable pour $\theta=\pi$ *puisque l'on connaît* $\displaystyle\sum_{n=0}^{+\infty}\frac{(-1)^n}{n+1} = \ln 2$ *(on peut aussi reprendre le calcul précédent), et pour* $\theta \in]\pi\,;2\pi[$ *on remarque que* $S(\theta) = \overline{S(2\pi-\theta)}$ *avec* $2\pi-\theta \in]0\,;\pi[$, *ce qui permet de montrer que la formule* $(*)$ *reste vraie pour* $0<\theta<2\pi$.

■ Comment estimer les sommes partielles (ou les restes) ?

Pour estimer la vitesse de convergence d'une série, ou pour essayer d'obtenir certains équivalents, il est parfois important de pouvoir encadrer les sommes partielles, ou le reste dans le cas d'une série convergente.

SF11.17 Comparer avec une série connue

Exemple

Pour tout $x>0$ on pose $f(x) = \displaystyle\sum_{n=1}^{+\infty} e^{-n^2x}$.

Justifier l'existence de $f(x)$ et en donner un équivalent quand $x \to +\infty$.

- *L'existence de $f(x)$ se démontre classiquement en utilisant la règle de comparaison* **[S11.6]**.
  En effet, pour $x > 0$ on a $\lim\limits_{n\to+\infty} n^2 \mathrm{e}^{-n^2x} = 0$ par les croissances comparées, donc $\mathrm{e}^{-n^2x} \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$, et puisque la série à termes réels positifs $\sum \frac{1}{n^2}$ est convergente, il en est de même de celle définissant $f(x)$.
- *On a pour tout $x > 0$:*
$$f(x) = \mathrm{e}^{-x} + \sum_{n=2}^{+\infty} \mathrm{e}^{-n^2x} = \mathrm{e}^{-x} + \mathrm{e}^{-4x} \sum_{n=2}^{+\infty} \mathrm{e}^{-(n^2-4)x}.$$
  Or pour $n \geqslant 2$, $n^2 - 4 \geqslant n - 2$ donc :
$$\sum_{n=2}^{+\infty} \mathrm{e}^{-(n^2-4)x} \leqslant \sum_{n=2}^{+\infty} \mathrm{e}^{-(n-2)x} = \frac{1}{1-\mathrm{e}^{-x}}$$
  (somme d'une série géométrique de raison e^{-x}).
  On aura donc, pour tout $x > 0$:
$$0 \leqslant f(x) - \mathrm{e}^{-x} \leqslant \frac{\mathrm{e}^{-4x}}{1-\mathrm{e}^{-x}}$$
  et puisque $\frac{\mathrm{e}^{-4x}}{1-\mathrm{e}^{-x}} \underset{x\to+\infty}{=} o(\mathrm{e}^{-x})$, on en déduit : $f(x) \underset{x\to+\infty}{\sim} \mathrm{e}^{-x}$.

SF11.18 Utiliser la méthode de comparaison à une intégrale

Nous avons déjà eu l'occasion d'utiliser cette méthode, qui est vraiment très importante dans l'étude des séries, convergentes ou non.

Exemple

Déterminer la partie entière de $\sum\limits_{n=1}^{10^9} \frac{1}{n^{2/3}}$.

Puisque la fonction $t \mapsto \frac{1}{t^{2/3}}$ est strictement décroissante sue $\mathbb{R}_+^*$, on a :
$$\forall\, n \geqslant 2,\ \int_n^{n+1} \frac{\mathrm{d}t}{t^{2/3}} < \frac{1}{n^{2/3}} < \int_{n-1}^{n} \frac{\mathrm{d}t}{t^{2/3}}.$$
Donc en additionnant ces inégalités pour n variant de 2 à 10^9 on obtient :
$$\int_2^{10^9+1} \frac{\mathrm{d}t}{t^{2/3}} < \sum_{n=2}^{10^9} \frac{1}{n^{2/3}} < \int_1^{10^9} \frac{\mathrm{d}t}{t^{2/3}}.$$
Une primitive de $t \mapsto t^{-2/3}$ étant $t \mapsto 3t^{1/3}$, on obtient alors :
$$3\left(\sqrt[3]{10^9+1} - \sqrt[3]{2}\right) < \sum_{n=2}^{10^9} \frac{1}{n^{2/3}} < 2997.$$
Le terme de gauche est plus grand que $3\left(1000 - \sqrt[3]{2}\right) \geqslant 2996$ (car $3\sqrt[3]{2} = \sqrt[3]{54} \leqslant \sqrt[3]{64} = 4$), donc, en ayant ajouté 1 :
$$2997 < \sum_{n=1}^{10^9} \frac{1}{n^{2/3}} < 2998,$$

et enfin : $\left\lfloor \sum_{n=1}^{10^9} \frac{1}{n^{2/3}} \right\rfloor = 2997.$

✓ *Pour pouvoir calculer la partie entière, il était ici indispensable de considérer des inégalités strictes, ce qui est en fait assez rare.*

SF11.19 Utiliser les résultats spécifiques aux séries alternées

Dans le critère spécial sur les séries alternées, il ne faut pas retenir que la condition de convergence, mais aussi les encadrements des sommes partielles et du reste.

Exemple

Pour tout $n \in \mathbb{N}$, on pose : $R_n = \sum_{k=n+1}^{+\infty} \frac{(-1)^k}{k}$.

Justifier l'existence de R_n et étudier la convergence de la série $\sum_{n\in\mathbb{N}} R_n$.

- *La série de terme général* $\frac{(-1)^k}{k}$ *vérifiant le critère spécial des séries alternées est convergente, ce qui justifie l'existence de* R_n.
- *Ce critère précise en outre que* R_n*, reste d'indice* n*, est du signe de* $(-1)^{n+1}$ *et que* $|R_n| \leqslant \frac{1}{n+1}$.

  Ainsi, la suite (R_n) *est alternée est converge vers* 0*. De plus :*

$$|R_n| = (-1)^{n+1} R_n = \left(\frac{1}{n+1} - \frac{1}{n+2}\right) + \left(\frac{1}{n+3} - \frac{1}{n+4}\right) + \cdots$$
$$= \frac{1}{(n+1)(n+2)} + \frac{1}{(n+3)(n+4)} + \cdots$$

  Sous cette forme, il est clair que $|R_{n+1}| \leqslant |R_n|$*. La série* $\sum R_n$ *vérifie donc le critère spécial des séries alternées, donc est convergente.*

SF11.20 Utiliser la sommation des relations de comparaison

Les résultats de **[S11.12]** permettent d'estimer les sommes partielles d'une série divergente ou les restes d'une série convergente par comparaison à une autre série.

Exemple

Montrer que pour $\alpha < 1$: $\sum_{k=1}^{n} \frac{1}{k^\alpha} \underset{n\to+\infty}{\sim} \frac{n^{1-\alpha}}{1-\alpha}$.

On obtient, par un développement limité rapide :

$$(n+1)^{1-\alpha} - n^{1-\alpha} = n^{1-\alpha}\left(\left(1+\frac{1}{n}\right)^{1-\alpha} - 1\right) \underset{n\to+\infty}{\sim} \frac{1-\alpha}{n^\alpha}.$$

La série de terme général $\frac{1}{n^\alpha}$ *étant divergente à termes positifs, le théorème de comparaison des sommes partielles des séries divergentes donne :*

$$\sum_{k=1}^{n} \frac{1}{k^\alpha} \underset{n\to+\infty}{\sim} \frac{1}{1-\alpha} \sum_{k=1}^{n} \left((k+1)^{1-\alpha} - k^{1-\alpha}\right) = \frac{(n+1)^{1-\alpha} - 1}{1-\alpha} \underset{n\to+\infty}{\sim} \frac{n^{1-\alpha}}{1-\alpha}.$$

✓ Pour obtenir cet équivalent classique, on pouvait aussi utiliser la comparaison avec une intégrale, comme en **[SF11.18]**.

Exemple

Pour $n \in \mathbb{N}^*$ on pose $u_n = \prod_{k=1}^{n} \left(1 + \frac{k}{n}\right)$. Trouver un équivalent de u_n quand $n \to +\infty$ à l'aide de la formule de Stirling puis montrer que :

$$u_n = \mathrm{e}^{-n} 2^{2n+1/2} \left(1 - \frac{1}{24n} + o\left(\frac{1}{n}\right)\right).$$

- *On peut écrire* $u_n = \dfrac{\prod_{k=1}^{n}(n+k)}{n^n} = \dfrac{(2n)!}{n!n^n}$. *Par la formule de Stirling on a alors l'équivalent :*

$$u_n \underset{n\to+\infty}{\sim} \frac{\left(\frac{2n}{\mathrm{e}}\right)^{2n}\sqrt{4\pi n}}{n^n\left(\frac{n}{\mathrm{e}}\right)^n\sqrt{2\pi n}} = \mathrm{e}^{-n}2^{2n+1/2}.$$

- *Posons* $v_n = u_n \mathrm{e}^n 2^{-2n-1/2}$ *pour* $n \in \mathbb{N}^*$. *Pour* $n \geqslant 2$ *on a :*

$$\frac{v_n}{v_{n-1}} = \frac{\mathrm{e}}{4}\frac{u_n}{u_{n-1}} = \frac{\mathrm{e}}{4}\frac{(2n)(2n-1)(n-1)^{n-1}}{nn^n} = \frac{\mathrm{e}}{2}\frac{2n-1}{n}\left(\frac{n-1}{n}\right)^{n-1}.$$

On effectue un développement limité :

$$\begin{aligned}\ln v_n - \ln v_{n-1} &= 1 + \ln\left(1 - \frac{1}{2n}\right) + (n-1)\ln\left(1 - \frac{1}{n}\right)\\ &= 1 - \frac{1}{2n} - \frac{1}{8n^2} + o\left(\frac{1}{n^2}\right) + (n-1)\left(-\frac{1}{n} - \frac{1}{2n^2} - \frac{1}{3n^3} + o\left(\frac{1}{n^3}\right)\right)\\ &= \frac{1}{24n^2} + o\left(\frac{1}{n^2}\right).\end{aligned}$$

La série de terme général $\frac{1}{24n^2}$ *étant une série à termes positifs convergente, le théorème de comparaison pour les restes des séries convergentes donne :*

$$\sum_{k=n+1}^{+\infty}\left(\ln v_k - \ln v_{k-1}\right) \underset{n\to+\infty}{\sim} \sum_{k=n+1}^{+\infty}\frac{1}{24k^2}.$$

Or $\lim_{k\to+\infty} \ln v_k = 0$ *donc* $\sum_{k=n+1}^{+\infty}\left(\ln v_k - \ln v_{k-1}\right) = -\ln v_n$. *D'autre part,* $\frac{1}{k^2} \underset{k\to+\infty}{\sim} \frac{1}{k(k-1)}$ *donc le même théorème donne :*

$$\sum_{k=n+1}^{+\infty}\frac{1}{k^2} \underset{n\to+\infty}{\sim} \sum_{k=n+1}^{+\infty}\frac{1}{k(k-1)} = \sum_{k=n+1}^{+\infty}\left(\frac{1}{k-1} - \frac{1}{k}\right) = \frac{1}{n},$$

donc finalement : $\ln v_n \sim -\dfrac{1}{24n}$.

On en déduit $v_n = \mathrm{e}^{-\frac{1}{24n} + o\left(\frac{1}{n}\right)} = 1 - \dfrac{1}{24n} + o\left(\dfrac{1}{n}\right)$, *ce qui donne le résultat demandé.*

■ Comment étudier la sommabilité d'une famille ?

Nous étudions dans ce paragraphe des familles $(u_i)_{i\in I}$ de nombres réels ou complexes indicées par un ensemble dénombrable I.

SF11.21 Majorer les sommes partielles

Nous appelons ici, de façon un peu abusive, « sommes partielles » d'une famille $(u_i)_{i\in I}$ les sommes $\displaystyle\sum_{i\in J} u_i$ où J est une partie finie de I.

Pour montrer la sommabilité d'une telle famille, on peut se contenter de démontrer que les sommes $\displaystyle\sum_{i\in J_n} |u_i|$ où (J_n) est une suite exhaustive de parties finies de I, sont majorées.

Exemple

Étudier la sommabilité (et calculer la somme éventuelle) de la famille $(u_n)_{n\in\mathbb{Z}}$ où :

$$\forall\, n\in\mathbb{Z},\ u_n = \frac{(-1)^n}{(|n|+1)^2}\cdot$$

- *On applique le critère mentionné ci-dessus à la suite exhaustive $(J_n)_{n\in\mathbb{N}}$ de parties de $\mathbb{Z}$ données par :*

$$\forall\, n\in\mathbb{N},\ J_n = [\![-n\,;n]\!]\,.$$

Or pour tout $n\in\mathbb{N}$:

$$\sum_{k=-n}^{n} |u_k| = \sum_{k=-n}^{n} \frac{1}{(|k|+1)^2} \leqslant 2\sum_{k=0}^{n} \frac{1}{(k+1)^2} \leqslant 2\sum_{k=0}^{+\infty} \frac{1}{(k+1)^2} = \frac{\pi^2}{3}\,,$$

donc la famille est sommable.

- *On a alors :*

$$\begin{aligned} s_{J_{2n}} = \sum_{k=-2n}^{2n} u_k &= -1 + 2\sum_{k=0}^{2n} \frac{(-1)^k}{(k+1)^2} \\ &= -1 + 2\left(\sum_{k=0}^{2n} \frac{1}{(k+1)^2} - 2\sum_{p=0}^{n-1} \frac{1}{(2p+2)^2}\right) \\ &= -1 + 2\sum_{k=0}^{2n} \frac{1}{(k+1)^2} - \sum_{p=0}^{n-1} \frac{1}{(p+1)^2}\,, \end{aligned}$$

et la somme totale de la famille est :

$$s = \lim_{n\to+\infty} s_{J_{2n}} = -1 + 2\sum_{k=0}^{+\infty} \frac{1}{(k+1)^2} - \sum_{p=0}^{\infty} \frac{1}{(p+1)^2} = \frac{\pi^2}{6} - 1\,.$$

✓ *On pouvait aussi utiliser ici le théorème de sommation par paquets, en écrivant $\mathbb{Z} = (-\mathbb{N})\cup\mathbb{N}^*$.*

SF11.22 Utiliser une majoration de $|u_i|$

Exemple

Soient x et y deux nombres réels strictement supérieurs à 1.
Montrer que la famille $\left(\dfrac{1}{x^n+y^m}\right)_{(n,m)\in\mathbb{N}^2}$ est sommable.

On utilise l'inégalité bien connue :

$$\forall\,(a,b)\in\mathbb{R}^2,\ 2ab\leqslant a^2+b^2\,,$$

qui permet d'écrire :

$$\forall\,(n,m)\in\mathbb{N}^2,\ x^n+y^m\geqslant 2x^{n/2}y^{m/2}\,,$$

puis :

$$\frac{1}{x^n+y^m}\leqslant\frac{1}{2}\left(x^{-1/2}\right)^n\left(y^{-1/2}\right)^m.$$

Or les séries à termes positifs $\displaystyle\sum_{n\in\mathbb{N}}\left(x^{-1/2}\right)^n$ *et* $\displaystyle\sum_{m\in\mathbb{N}}\left(y^{-1/2}\right)^m$ *sont convergentes (car géométriques de raison* <1*), donc la famille* $\left(\left(x^{-1/2}\right)^n\left(y^{-1/2}\right)^m\right)_{(n,m)\in\mathbb{N}^2}$ *est sommable (c'est un corollaire du théorème de Fubini).*

Par comparaison, la série proposée est sommable.

SF11.23 Utiliser un regroupement par paquets

Le théorème de sommation par paquets permet non seulement de montrer la sommabilité d'une famille, mais aussi de calculer sa somme.

Exemple

Étudier la sommabilité de la suite double $(a_{p,q})_{(p,q)\in\mathbb{N}^2}$ définie par :

$$\forall\,(p,q)\in\mathbb{N}^2,\ a_{p,q}=\frac{1}{(p+q+1)^\alpha}\quad(\alpha\in\mathbb{R}).$$

On considère naturellement la partition de $\mathbb{N}^2$ *formée des ensembles :*

$$\forall\,n\in\mathbb{N},\ \Delta_n=\left\{(p,q)\in\mathbb{N}^2\;\middle|\;p+q=n\right\},$$

et on applique le théorème de sommation par paquets :

- *pour tout* $n\in\mathbb{N}$*, la famille* $(a_{p,q})_{(p,q)\in\Delta_n}$ *est sommable : c'est immédiat puisque l'ensemble* Δ_n *est de cardinal fini ;*
- *notons alors* $s_n=\displaystyle\sum_{(p,q)\in\Delta_n}a_{(p,q)}$*. On a :*

$$s_n=\sum_{p+q=n}\frac{1}{(p+q+1)^\alpha}=(n+1)\cdot\frac{1}{(n+1)^\alpha}=\frac{1}{(n+1)^{\alpha-1}}\cdot$$

La famille est donc sommable si et seulement si la série $\sum s_n$ *est (absolument) convergente, c'est-à-dire si* $\alpha>2$*.*

De plus, dans ce cas, on a :

$$\sum_{(p,q)\in\mathbb{N}^2}a_{p,q}=\sum_{n=0}^{+\infty}\frac{1}{(n+1)^{\alpha-1}}\cdot$$

SF11.24 Utiliser le théorème de Fubini

Les suites doubles sont fréquentes, et le théorème de Fubini est l'un des moyens privilégiés pour montrer leur sommabilité et calculer leur somme.

Exemple

1) Existence et calcul de : $\displaystyle\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}^*} \frac{1}{(p+q^2)(p+q^2+1)}$.

2) En déduire la valeur de la somme : $\displaystyle\sum_{n=1}^{+\infty} \frac{\lfloor\sqrt{n}\rfloor}{n(n+1)}$.

1) *Pour* $(p,q)\in\mathbb{N}\times\mathbb{N}^*$, *notons* $u_{p,q}=\dfrac{1}{(p+q^2)(p+q^2+1)}$.

Par une décomposition en éléments simples on obtient :

$$u_{p,q}=\frac{1}{p+q^2}-\frac{1}{p+q^2+1}.$$

Pour tout $q\in\mathbb{N}^*$, *la série* $\displaystyle\sum_{p\in\mathbb{N}} u_{p,q}$ *est à termes positifs, converge, et a pour somme, par télescopage :* $\displaystyle\sum_{p=0}^{+\infty} u_{p,q}=\frac{1}{q^2}$.

Puisque la série $\displaystyle\sum_{q\in\mathbb{N}^*}\frac{1}{q^2}$ *converge, le théorème de Fubini affirme que la suite double* $(u_{p,q})$ *est sommable et que :*

$$\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}^*} u_{p,q}=\sum_{q=1}^{+\infty}\frac{1}{q^2}=\frac{\pi^2}{6}.$$

2) *Considérons la partition de* $\mathbb{N}\times\mathbb{N}^*$ *formée des ensembles :*

$$I_n=\left\{(p,q)\in\mathbb{N}\times\mathbb{N}^* \;\middle|\; p+q^2=n\right\} \quad \text{avec } n\in\mathbb{N}^*.$$

Puisque la famille $(u_{p,q})$ *est sommable, le théorème de sommation par paquets affirme que :*

$$\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}^*} u_{p,q}=\sum_{n=1}^{+\infty}\left(\sum_{(p,q)\in I_n} u_{p,q}\right).$$

Or : $\displaystyle\sum_{(p,q)\in I_n} u_{p,q}=\sum_{p+q^2=n}\frac{1}{(p+q^2)(p+q^2+1)}=\frac{\mathrm{Card}(I_n)}{n(n+1)}$,

et le cardinal de I_n *est exactement égal au nombre d'entiers* $q\in\mathbb{N}^*$ *tels que* $q^2\leqslant n$, *c'est-à-dire à* $\lfloor\sqrt{n}\rfloor$. *On a donc finalement :*

$$\sum_{n=1}^{+\infty}\frac{\lfloor\sqrt{n}\rfloor}{n(n+1)}=\sum_{(p,q)\in\mathbb{N}\times\mathbb{N}^*} u_{p,q}=\frac{\pi^2}{6}.$$

Le théorème de Fubini peut également permettre de démontrer certaines égalités en faisant apparaître des suites doubles.

Exemple

Soit z un nombre complexe tel que $|z|<1$. Montrer que :

$$\sum_{n=1}^{+\infty}\frac{z^n}{1-z^{2n}}=\sum_{n=1}^{+\infty}\frac{z^{2n-1}}{1-z^{2n-1}}\cdot$$

- *Déjà, les deux séries proposées sont absolument convergentes puisque :*

$$\left|\frac{z^n}{1-z^{2n}}\right|\underset{n\to+\infty}{\sim}|z|^n \quad \text{et} \quad \left|\frac{z^{2n-1}}{1-z^{2n-1}}\right|\underset{n\to+\infty}{\sim}|z|^{2n-1},$$

avec $|z|<1$ (comparaison à une série géométrique).

- *Pour tout $n\in\mathbb{N}^*$ on a :* $\dfrac{z^n}{1-z^{2n}}=z^n\sum_{k=0}^{+\infty}\left(z^{2n}\right)^k=\sum_{k=0}^{+\infty}z^{n(2k+1)}$.

Posons alors pour $(n,k)\in\mathbb{N}^*\times\mathbb{N}$, $u_{n,k}=z^{n(2k+1)}$. On applique à cette suite double le théorème de Fubini :

– *pour tout $n\in\mathbb{N}^*$, la série $\sum_{k\in\mathbb{N}}u_{n,k}$ est absolument convergente, puisque $\left|z^{2n}\right|<1$, et elle a pour somme :*

$$s_n=\sum_{k=0}^{+\infty}|u_{n,k}|=|z|^n\sum_{k=0}^{+\infty}\left(|z|^{2n}\right)^k=\frac{|z|^n}{1-|z|^{2n}};$$

– *la série $\sum_{n\in\mathbb{N}^*}s_n$ est convergente puisque $s_n\underset{n\to+\infty}{\sim}|z|^n$.*

Les hypothèses du théorème sont donc bien vérifiées ; on en conclut que la suite double $(u_{n,k})_{(n,k)\in\mathbb{N}^*\times\mathbb{N}}$ est sommable et que :

$$\begin{aligned}\sum_{(n,k)\in\mathbb{N}^*\times\mathbb{N}}u_{n,k}&=\sum_{k=0}^{+\infty}\left(\sum_{n=1}^{+\infty}\left(z^{2k+1}\right)^n\right)=\sum_{k=0}^{+\infty}\frac{z^{2k+1}}{1-z^{2k+1}}\\&=\sum_{n=1}^{+\infty}\left(\sum_{k=0}^{+\infty}z^n\left(z^{2n}\right)^k\right)=\sum_{n=1}^{+\infty}\frac{z^n}{1-z^{2n}},\end{aligned}$$

ce qui est le résultat demandé.

Enfin, le théorème de Fubini peut permettre d'intervertir des sommes pour obtenir par exemple des développements en série entière ou des développements limités.

Exemple

Soit a un réel tel que $|a|<1$. On considère la fonction :

$$S\colon x\in\,]-1\,;1[\longmapsto\sum_{n=1}^{+\infty}\frac{a^n}{x+n}\cdot$$

Montrer que S est développable en série entière sur $]-1\,;1[$.

- *Chaque fonction $x\mapsto\dfrac{a^n}{x+n}$ est bien définie sur $]-1\,;1[$ puisque $n\geqslant 1$.*
De plus, la série donnée est absolument convergente puisque, pour $x\in\,]-1\,;1[$, $\left|\dfrac{a^n}{x+n}\right|\underset{n\to+\infty}{\sim}\dfrac{|a|^n}{n}$ qui est le terme général d'une série convergente (de somme $-\ln(1-|a|)$).
Donc S est bien définie sur $]-1\,;1[$.

- *Pour tout $x \in]-1\,;1[$ et tout $n \in \mathbb{N}^*$, on a :*

$$\frac{1}{x+n} = \frac{1}{n\left(1+\frac{x}{n}\right)} = \frac{1}{n}\sum_{k=0}^{+\infty}\left(-\frac{x}{n}\right)^k = \sum_{k=0}^{+\infty}(-1)^k\frac{x^k}{n^{k+1}}\,,$$

donc : $S(x) = \sum\limits_{(n,k)\in\mathbb{N}^*\times\mathbb{N}} u_{n,k}$ *avec* $u_{n,k} = (-1)^k\dfrac{a^n x^k}{n^{k+1}}$.

Nous allons appliquer le théorème de Fubini à cette suite double ($x \in]-1\,;1[$ étant fixé) :

– *pour tout $n \in \mathbb{N}^*$, la série $\sum\limits_{k=0}^{+\infty}|u_{n,k}|$ converge, et on peut calculer, par des calculs similaires aux précédents :*

$$v_n = \sum_{k=0}^{+\infty}|u_{n,k}| = |a|^n\sum_{k=0}^{+\infty}\frac{|x|^k}{n^{k+1}} = \frac{|a|^n}{n}\frac{1}{1-\frac{|x|}{n}} = \frac{|a|^n}{n-|x|}\,.$$

– *la série de terme général v_n est convergente, puisque* $v_n \underset{n\to+\infty}{\sim} \dfrac{|a|^n}{n}$.

On en conclut que la famille $(u_{n,k})_{(n,k)\in\mathbb{N}^*\times\mathbb{N}}$ est sommable, et l'on peut intervertir l'ordre des sommations, ce qui donne :

$$\forall\, x \in]-1\,;1[,\ S(x) = \sum_{k=0}^{+\infty}(-1)^k\left(\sum_{n=1}^{+\infty}\frac{a^n}{n^{k+1}}\right)x^k\,.$$

Cela prouve que S est développable en série entière sur $]-1\,;1[$, le coefficient de x^k dans ce développement étant égal à $(-1)^k\sum\limits_{n=1}^{+\infty}\dfrac{a^n}{n^{k+1}}$.

À vous de jouer !

Exercice 11.1 **SF11.4** **SF11.10**

Étudier la nature de la série de terme général $u_n = \ln\left(1+\sin\left(\frac{(-1)^n}{n^\alpha}\right)\right)$ selon les valeurs du réel $\alpha > 0$.

Exercice 11.2 **SF11.10**

Nature des séries de terme général u_n et $(-1)^n u_n$ où :

$$u_n = n + \frac{1}{2} - \frac{1}{\ln(n^2+n+1) - \ln(n^2+1)}\,.$$

Exercice 11.3 **SF11.11**

Soit $(a_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{R}_+$. On définit la suite $(u_n)_{n\in\mathbb{N}}$ par :

$$u_0 > 0 \quad \text{et} \quad \forall\, n \in \mathbb{N},\ u_{n+1} = u_n + \frac{a_n}{u_n}\,.$$

Montrer que : la suite (u_n) converge $\iff$ la série $\sum a_n$ converge.

Exercice 11.4 **SF11.1** **SF11.2** **SF11.11** **SF11.18**

$\sum u_n$ désigne une série à termes strictement positifs, divergente, et l'on note : $s_n = \sum_{k=0}^{n} u_k$.

1) Montrer que la série de terme général $\dfrac{u_n}{s_n}$ diverge (on pourra considérer la série de terme général $\ln(s_n) - \ln(s_{n-1})$).

2) Étudier, selon les valeurs de α, la nature de la série de terme général $\dfrac{u_n}{s_n^\alpha}$ (dans le cas $\alpha > 1$, on pourra comparer à la série de terme général $\displaystyle\int_{s_{n-1}}^{s_n} \frac{\mathrm{d}t}{t^\alpha}$).

Exercice 11.5 **SF11.2** **SF11.18**

Notons $f\colon x \mapsto \dfrac{\ln x}{x}$ pour $x > 0$.

1) Justifier la convergence de la série $\displaystyle\sum_{n\geqslant 1} (-1)^n f(n)$.

2) Calculer la différence : $\displaystyle 2\sum_{k=1}^{n} f(2k) - \sum_{k=1}^{2n} f(k)$. En déduire la valeur de $\displaystyle\sum_{n=1}^{+\infty} (-1)^n f(n)$ en fonction de la constante d'Euler.

Exercice 11.6 **SF11.1** **SF11.15**

Montrer que, si α et β sont deux réels strictement positifs tels que $\alpha + \beta \leqslant 1$, la série produit de Cauchy des séries $\displaystyle\sum_{n\geqslant 0} \frac{(-1)^n}{(n+1)^\alpha}$ et $\displaystyle\sum_{n\geqslant 0} \frac{(-1)^n}{(n+1)^\beta}$ est divergente.

Exercice 11.7 **SF11.20**

Pour $n \in \mathbb{N}^*$ on pose : $u_n = \dfrac{n!\mathrm{e}^n}{n^{n+1/2}}$. En considérant la série de terme général $\ln\left(\dfrac{u_n}{u_{n+1}}\right)$, établir le développement asymptotique :

$$n! \underset{n\to+\infty}{=} \left(\frac{n}{\mathrm{e}}\right)^n \sqrt{2\pi n}\left(1 + \frac{1}{12n} + \frac{1}{288n^2} + O\left(\frac{1}{n^3}\right)\right).$$

Exercice 11.8 **SF11.23**

Justifier l'existence et calculer : $\displaystyle\sum_{(p,q)\in\mathbb{N}^2} \frac{1}{p!q!(p+q+1)}$.

Exercice 11.9 **SF11.23** **SF11.24**

Pour tout entier $q \geqslant 2$ on note : $\zeta(q) = \displaystyle\sum_{p=1}^{+\infty} \frac{1}{p^q}$.

Montrer que pour tout $z \in \mathbb{C}$ tel que $|z| < 2$, on a :

$$\sum_{p=2}^{+\infty} \frac{z^2}{p(p-z)} = \sum_{q=2}^{+\infty} \big(\zeta(q) - 1\big)\, z^q.$$

Exercice 11.10 **SF11.21** **SF11.23** **SF11.24**

1) Montrer que la suite double $\left(\dfrac{1}{p^2q^2}\right)_{(p,q)\in(\mathbb{N}^*)^2}$ est sommable et calculer sa somme.

2) En déduire la valeur de : $\displaystyle\sum_{\substack{(p,q)\in(\mathbb{N}^*)^2 \\ p\wedge q=1}} \frac{1}{p^2q^2}$.

On rappelle : $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ et $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$.

Exercice 11.11 **SF11.24**

Montrer que pour tout $z \in \mathbb{C}$ tel que $|z| < 1$ on a :

$$\sum_{n=1}^{+\infty} \frac{z^n}{1-z^n} = \sum_{n=1}^{+\infty} d(n) z^n\,,$$

où $d(n)$ désigne le nombre de diviseurs de n.

Exercice 11.12 **SF11.24**

Déterminer un développement limité à tout ordre selon les puissances de $\dfrac{1}{n}$ de $u_n = \displaystyle\int_0^1 \frac{\ln t}{1+t^n}\,\mathrm{d}t$ lorsque $n \to +\infty$.

■ Comment étudier la convergence simple d'une suite de fonctions ?

SF12.1 Étudier la convergence d'une suite

Étudier la convergence simple d'une suite de fonctions (f_n), c'est, par définition, étudier la suite $(f_n(x))$. Cependant, pour l'étude de cette suite quand $n \to +\infty$, il ne faut pas perdre de vue que la variable x joue le rôle d'un paramètre et que l'existence et la valeur de la limite peuvent amener à distinguer plusieurs cas selon le domaine de variations de x.

De plus, il ne faut pas oublier que le domaine de définition de la limite simple f d'une suite de fonctions (f_n) est nécessairement inclus dans l'intersection des domaines de définition des f_n. Avant de se lancer dans le calcul de $\lim\limits_{n\to+\infty} f_n(x)$, il faut donc vérifier que les expressions $f_n(x)$ ont bien un sens.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies sur $]-1\,;+\infty[$ par :

$$\forall\, x > -1,\ f_n(x) = \frac{x^n}{1+x^n}\cdot$$

On commence tout de suite par remarquer que :

- *toutes les f_n sont bien définies sur $]-1\,;+\infty[$;*
- *la présence du terme x^n impose de discuter selon la position de $|x|$ par rapport à 1.*

- *Si $|x|<1$, $\lim\limits_{n\to+\infty} x^n = 0$ donc $\lim\limits_{n\to+\infty} f_n(x) = 0$.*
- *Si $x = 1$, $f_n(1) = \frac{1}{2}$ pour tout n donc $\lim\limits_{n\to+\infty} f_n(1) = \frac{1}{2}$.*
- *Si $x > 1$, $\lim\limits_{n\to+\infty} x^n = +\infty$ donc $1 + x^n \underset{x\to+\infty}{\sim} x^n$ et $\lim\limits_{n\to+\infty} f_n(x) = 1$.*

En résumé, la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur $]-1\,;+\infty[$ vers la fonction $f : x \mapsto \begin{cases} 0 & \text{si } x \in]-1\,;0[\\ \frac{1}{2} & \text{si } x = 1 \\ 1 & \text{si } x > 1\,. \end{cases}$

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ définies sur $\mathbb{R}$ par :

$$f_n(x) = \begin{cases} \sin(\pi n x) & \text{si } |x| < \frac{1}{n} \\ 0 & \text{sinon}\,. \end{cases}$$

✓ Il s'agit ici d'un exemple où la définition de f_n varie selon les intervalles où se situe x, ces intervalles dépendant de n.

La manière d'aborder cette situation est toujours la même : conformément à la définition de la convergence *simple*, il faut fixer un x et examiner dans quel intervalle il se situe, au moins pour n assez grand.

x étant fixé, pour déterminer dans quel intervalle il se situe on est amené à distinguer deux cas :

- *si $x = 0$, alors pour tout $n \in \mathbb{N}^*$ on a $|x| < \frac{1}{n}$ donc $f_n(0) = \sin 0 = 0$;*
- *si $x \neq 0$, alors on aura $|x| \geqslant \frac{1}{n}$ à partir d'un certain rang (plus précisément pour $n \geqslant \frac{1}{|x|}$) donc $f_n(x) = 0$ à partir d'un certain rang et par suite $\lim\limits_{n\to+\infty} f_n(x) = 0$.*

En conclusion, la suite (f_n) converge simplement sur $\mathbb{R}$ vers la fonction nulle. On notera que la convergence n'est pas uniforme sur $\mathbb{R}$ puisque de façon claire $\|f_n\|_\infty^{\mathbb{R}} = 1$.

SF12.2 Prouver directement la convergence uniforme

Dans certains cas, il est possible, si l'on a l'intuition de la limite ou si celle-ci est donnée par l'énoncé, de prouver directement la convergence uniforme.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ définies sur $\mathbb{R}_+$ par :

$$\forall\, x \geqslant 0,\ f_n(x) = \frac{x}{n(1+x^n)}\cdot$$

Pour tout $n \geqslant 1$, f_n est bien définie sur $\mathbb{R}_+$.
Pour $x \in [0\,;1]$ on a $0 \leqslant f_n(x) \leqslant \frac{x}{n} \leqslant \frac{1}{n}$ et pour $x \geqslant 1$, on a :

$$0 \leqslant f_n(x) \leqslant \frac{x}{nx^n} = \frac{1}{nx^{n-1}} \leqslant \frac{1}{n}\cdot$$

Donc $\|f_n\|_\infty^{\mathbb{R}_+} \leqslant \frac{1}{n} \underset{n\to+\infty}{\longrightarrow} 0$, ce qui prouve que la suite (f_n) converge uniformément sur $\mathbb{R}_+$ vers la fonction nulle.

■ Comment étudier la convergence uniforme d'une suite de fonctions ?

On considère ici une suite de fonctions $(f_n)_{n\in\mathbb{N}}$, définies sur un intervalle I de $\mathbb{R}$. On suppose que l'on a déjà étudié la convergence simple de cette suite, et que l'on a déterminé la limite simple f, définie sur I (quitte à réduire l'intervalle I s'il n'y a pas convergence simple partout), et l'on s'intéresse ensuite à la convergence uniforme éventuelle.

✓ On parle de la convergence uniforme de la suite (f_n) vers f **sur I**. Cela n'a *aucun sens* de dire : « la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f » sans préciser sur quel ensemble, conformément à la définition **[S12.1]**.

De même, si les fonctions f_n dépendent de plusieurs variables, comme cela

peut être le cas lors de l'étude d'intégrales à paramètres, il faut préciser **par rapport** à quelle variable on considère la convergence uniforme.

SF12.3 Revenir à la définition avec le sup

Cette méthode consiste simplement à démontrer que $\lim\limits_{n\to+\infty} \|f_n - f\|_\infty^I = 0$.
Pour cela, on forme la différence $f_n - f$; ensuite, soit on calcule explicitement la borne supérieure (à l'aide d'une étude de fonctions par exemple), soit, si ce calcul paraît trop compliqué, on essaye de la majorer par une quantité qui tend vers 0.

Exemple

Étudier la suite de fonctions $f_n : x \mapsto \sin x \cos^n x$ sur $\left[0\,;\frac{\pi}{2}\right]$.

On commence par étudier rapidement la convergence simple :

– *pour tout* n, $f_n(0) = 0$ *donc* $\lim\limits_{n\to+\infty} f_n(0) = 0$;

– *si* $x \in]0\,;1]$, $|\cos x| < 1$ *donc* $\lim\limits_{n\to+\infty} \cos^n x = 0$ *puis* $\lim\limits_{n\to+\infty} f_n(x) = 0$.

En conclusion, la suite $(f_n)_{n\in\mathbb{N}}$ *converge simplement sur* $[0\,;1]$ *vers la fonction nulle.*

Pour prouver la convergence uniforme, il faut donc montrer que $\lim\limits_{n\to+\infty} \|f_n\|_\infty^{[0;1]} = 0$. *Pour cela on va faire une étude rapide et former le tableau de variations de* f_n *pour* $n \geqslant 1$.

$\forall\, x \in [0\,;1]$, $f_n'(x) = \cos^{n-1} x \left(\cos^2 x - n\sin^2 x\right)$ *donc* f_n' *est, sur* $\left[0\,;\frac{\pi}{2}\right[$, *du signe de* $1 - n\tan^2 x$. *Sur cet intervalle,* f_n' *s'annule en* $\alpha_n = \operatorname{Arctan}\left(\frac{1}{\sqrt{n}}\right)$ *et on a donc facilement le tableau suivant :*

x	0		α_n		$\frac{\pi}{2}$
f_n	0	↗		↘	0

Ce tableau montre que $\|f_n\|_\infty = f_n(\alpha_n)$. *On pourrait bien sûr utiliser les relations :*

$$\forall\, x \geqslant 0,\ \cos(\operatorname{Arctan} x) = \frac{1}{\sqrt{1+x^2}} \quad \text{et} \quad \sin(\operatorname{Arctan} x) = \frac{x}{\sqrt{1+x^2}}$$

pour calculer explicitement $f_n(\alpha_n)$ *mais il est bien plus rapide de remarquer que, puisque* $0 \leqslant \cos x \leqslant 1$ *et* $0 \leqslant \sin x \leqslant x$ *sur* $\left[0\,;\frac{\pi}{2}\right]$, *on a* $0 \leqslant f_n(x) \leqslant x$, *et puisque* $0 \leqslant \operatorname{Arctan} x \leqslant x$ *pour* $x \geqslant 0$ *(position de la courbe de* Arctan *par rapport à sa tangente en* 0*), on aura* $f_n(\alpha_n) \leqslant \alpha_n \leqslant \frac{1}{\sqrt{n}}$.

Cela montre que $\lim\limits_{n\to+\infty} \|f_n\|_\infty = 0$, *donc que la suite* $(f_n)_{n\in\mathbb{N}}$ *converge uniformément vers la fonction nulle sur* $[0\,;1]$.

Exemple

Étudier la suite de fonctions $(f_n)_{n\geqslant 1}$ définies sur $\mathbb{R}$ par :

$$f_n(x) = \begin{cases} 0 & \text{si } 0 \leqslant x \leqslant 1/n \\ \dfrac{1}{n\ln\left(1-\frac{1}{nx}\right)} & \text{si } x<0 \text{ ou } x>1/n\,. \end{cases}$$

- *On commence déjà par remarquer que la définition a bien un sens (ne pas oublier ce point !). En effet, lorsque* $x \in]-\infty\,;0[\,\cup\,\left]\frac{1}{n}\,;+\infty\right[$, *on a* $1-\frac{1}{nx} = \frac{nx-1}{nx}$ *strictement positif et différent de* 1 *donc* $\ln\left(1-\frac{1}{nx}\right)$ *a un sens et ne s'annule pas.*
- *On étudie ensuite la convergence simple :*
  - *pour tout* $n \in \mathbb{N}^*$, $f_n(0) = 0$ *donc* $\lim\limits_{n\to+\infty} f_n(0) = 0$.
  - *si* $x \neq 0$, *soit* x *est négatif, soit il est strictement positif et l'on aura forcément* $x > \frac{1}{n}$ *pour* n *assez grand. Dans les deux cas, on aura, au moins à partir d'un certain rang,* $f_n(x) = \dfrac{1}{n\ln\left(1-\frac{1}{nx}\right)}$ *et puisque* $\ln\left(1-\frac{1}{nx}\right) \underset{n\to+\infty}{\sim} -\frac{1}{nx}$ *on a* $\lim\limits_{n\to+\infty} f_n(x) = -x$.

  En conclusion, la suite $(f_n)_{n\in\mathbb{N}}$ *converge simplement sur* $\mathbb{R}$ *vers la fonction* $f\colon x \mapsto -x$.
- *Pour la convergence uniforme, il faut étudier* $\|f_n - f\|_\infty^{\mathbb{R}}$. *Mais ici, l'étude de la fonction* $f_n - f$ *ne semble guère faisable. Il faut donc chercher à majorer* $|f_n(x) - f(x)|$ *directement, sans faire d'étude de fonction.*

  On remarque pour cela que, pour $x \in]-\infty\,;0[\cup\left]\frac{1}{n}\,;+\infty\right[$, *on peut écrire :*

$$f_n(x) - f(x) = \frac{1}{n\ln\left(1-\frac{1}{nx}\right)} + x = \frac{1}{n}\left(\frac{1}{\ln\left(1-\frac{1}{nx}\right)} + nx\right) = \frac{1}{n}\varphi(nx)$$

où $\varphi\colon \begin{cases}]-\infty\,;0[\cup]1\,;+\infty[& \longrightarrow \mathbb{R} \\ t & \longmapsto \dfrac{1}{\ln\left(1-\frac{1}{t}\right)} + t\,. \end{cases}$

Or la fonction φ *est continue sur* $]-\infty\,;0[\cup]1\,;+\infty[$ *par les théorèmes généraux ; deux calculs faciles donnent* $\lim\limits_{t\to 0^-}\varphi(t) = 0$ *et* $\lim\limits_{t\to 1^+}\varphi(t) = 1$, *donc* φ *se prolonge en une fonction continue sur* $]-\infty\,;0]\cup[1\,;+\infty[$.

Enfin, un développement limité montre que $1 + t\ln\left(1-\frac{1}{t}\right) \underset{t\to\pm\infty}{\sim} -\frac{1}{2t}$

donc $\lim\limits_{t\to\pm\infty}\varphi(t) = \frac{1}{2}$.

Il en résulte que φ *est bornée sur son domaine de définition, et donc, pour tout* $x \in]-\infty\,;0[\cup\left]\frac{1}{n}\,;+\infty\right[$, $|f_n(x) - f(x)| \leqslant \dfrac{\|\varphi\|_\infty}{n}$.

Puisque de plus, $|f_n(x) - f(x)| = |x| \leqslant \frac{1}{n}$ *lorsque* $x \in \left[0\,;\frac{1}{n}\right]$, *on a :*

$$\|f_n - f\|_\infty^{\mathbb{R}} \leqslant \frac{\max(1, \|\varphi\|_\infty)}{n}\,,$$

ce qui implique la convergence uniforme de la suite (f_n) vers f sur $\mathbb{R}$.

SF12.4 Revenir à la définition avec les ε

Traduite formellement, la propriété « la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur I » peut s'écrire :

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ \forall x \in I,\ |f_n(x) - f(x)| < \varepsilon\,.$$

✓ Il faut bien réfléchir ici à la place des quantificateurs $\forall$ et $\exists$ dans cette phrase : l'entier n_0 à partir duquel on a atteint la précision ε demandée ne dépend *que* de ε et non de la valeur de $x \in I$; c'est pour cette raison que l'on parle de convergence *uniforme*.

La manipulation des ε étant une chose délicate (même si nous estimons que c'est à la base de l'analyse), il ne faut revenir à la définition qu'en dernier recours, ou bien dans les problèmes théoriques.
Pour illustrer cette méthode, nous reprenons un exemple vu plus haut. Vous constaterez que bien que plus difficile à comprendre, cette méthode est ici plus rapide.

Exemple

Étudier la suite de fonctions $f_n : x \mapsto \sin x \cos^n x$ sur $\left[0\,;\frac{\pi}{2}\right]$.

Nous avons déjà vu que la suite (f_n) converge simplement sur $\left[0\,;\frac{\pi}{2}\right]$ vers la fonction nulle. Démontrons maintenant qu'il y a convergence uniforme sur cet intervalle.

L'idée est ici de partager l'intervalle $[0\,;1]$ en deux parties : l'une sur laquelle on aura $|\cos x| < 1$, puisqu'alors $\lim\limits_{n\to+\infty} (\cos x)^n = 0$, l'autre où l'on utilise $\lim\limits_{x\to 0} \sin x = 0$.

Soit donc un réel $\varepsilon > 0$ donné, et un réel α dans $\left]0\,;\frac{\pi}{2}\right]$ (que l'on précisera par la suite). Compte tenu des inégalités $|\sin x| \leqslant |x|$ et $|\cos x| \leqslant 1$ valables pour tout x réel, on a :

$$\forall x \in [0\,;\alpha],\ |f_n(x)| = |\sin x|\,|\cos x|^n \leqslant \alpha \quad (1)$$

D'autre part, la fonction cos *étant décroissante sur $\left[0\,;\frac{\pi}{2}\right]$, on aura*

$$\forall x \in \left[\alpha\,;\tfrac{\pi}{2}\right],\ |f_n(x)| \leqslant (\cos\alpha)^n \quad (2)$$

Or $|\cos\alpha| < 1$ donc $\lim\limits_{n\to+\infty} (\cos\alpha)^n = 0$ ce qui signifie qu'il existe un entier n_0 tel que, pour tout $n \geqslant n_0$, on a $|\cos\alpha|^n < \varepsilon$.

Finalement, en choisissant $\alpha \in\,]0\,;\varepsilon[$, la combinaison des inégalités (1) et (2) donne :

$$\forall n \geqslant n_0\ ,\ \forall x \in \left[0\,;\tfrac{\pi}{2}\right]\ ,\ |f_n(x)| < \varepsilon\,,$$

ce qui prouve la convergence uniforme de la suite (f_n) vers la fonction nulle sur $\left[0\,;\frac{\pi}{2}\right]$.

SF12.5 Étudier la convergence uniforme locale

Si la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ converge simplement mais pas uniformément vers f sur I, on peut alors essayer de trouver un intervalle $J \subset I$ sur lequel il y a convergence uniforme.
Cela peut être utile pour appliquer certains théorèmes que nous verrons ensuite et qui exigent la convergence uniforme, au moins locale, comme hypothèse.

Exemple

Étudier la convergence (simple, uniforme, uniforme locale) de la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies sur $\mathbb{R}_+^*$ par : $\forall\, x > 0,\ f_n(x) = \dfrac{nx^2\mathrm{e}^{-nx}}{(1-\mathrm{e}^{-x})^2}$.

- *(f_n) converge simplement vers la fonction nulle sur $\mathbb{R}_+^*$, puisque pour $x > 0$, $\lim\limits_{n\to+\infty} nx^2\mathrm{e}^{-nx} = 0$ d'après les croissances comparées.*
- *Il ne peut y avoir convergence uniforme sur $\mathbb{R}_+^*$ car alors par le théorème de la double limite on aurait $\lim\limits_{x\to0^+}\lim\limits_{n\to+\infty} f_n(x) = \lim\limits_{n\to+\infty}\lim\limits_{x\to0^+} f_n(x)$; or $\lim\limits_{x\to0^+} f_n(x) = n$ puisque $1-\mathrm{e}^{-x} \underset{x\to0}{\sim} x$, d'où la contradiction.*
- *Pour $a > 0$, sur $[a\,;+\infty[$, on a : $|f_n(x)| \leqslant \dfrac{nx^2\mathrm{e}^{-nx}}{(1-\mathrm{e}^{-a})^2}$. L'étude rapide de la fonction $x \mapsto x^2\mathrm{e}^{-nx}$ (maximum en $x = \dfrac{2}{n}$) montre que $nx^2\mathrm{e}^{-nx} \leqslant \dfrac{4}{n}\mathrm{e}^{-2}$ donc $\|f_n\|_\infty^{[a;+\infty[} \leqslant \dfrac{4\mathrm{e}^{-2}}{n(1-\mathrm{e}^{-a})^2}$ et $\|f_n\|_\infty^{[a;+\infty[}$ tend vers 0 quand $n \to +\infty$, ce qui donne la convergence uniforme sur $[a\,;+\infty[$.*
- *Autre solution : une majoration plus astucieuse est possible, qui évite l'étude d'une fonction auxiliaire.*
  La fonction $g\colon x \mapsto \dfrac{x^2}{(1-\mathrm{e}^{-x})^2}$ est continue sur $\mathbb{R}_+^*$, $\lim\limits_{x\to0^+} g(x) = 1$ et $\lim\limits_{x\to+\infty} g(x) = 0$ donc g est bornée sur $\mathbb{R}_+^*$.
  On en déduit, pour $x > a$, $|f_n(x)| \leqslant n\mathrm{e}^{-nx}\|g\|_\infty \leqslant n\mathrm{e}^{-na}\|g\|_\infty$ d'où la convergence uniforme sur $[a\,;+\infty[$.

■ Comment prouver la *non* convergence uniforme d'une suite de fonctions ?

SF12.6 Revenir à la définition

Pour prouver que la suite (f_n), qui converge déjà simplement vers f sur I, ne converge pas uniformément, il suffit de montrer, conformément à la définition, que $\|f_n - f\|_\infty$ ne tend pas vers 0. Pour cela on peut :

– calculer explicitement cette quantité ;

– la minorer par un terme ne tendant pas vers 0 lorsque $n \to +\infty$; pour cela on peut essayer de trouver une valeur $x_n \in I$ telle que $(f_n - f)(x_n)$ ne tende pas vers 0 lorsque $n \to +\infty$; puisque $\|f_n - f\|_\infty \geqslant |(f_n - f)(x_n)|$, on peut conclure.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ où : $\forall\, x \in \mathbb{R}$, $f_n(x) = \dfrac{nx}{1+n^2x^2}$.

Pour $x = 0$, $f_n(0) = 0$ pour tout n. Si $x \neq 0$, $f_n(x) \underset{n\to+\infty}{\sim} \dfrac{nx}{n^2x^2} = \dfrac{1}{nx}$ qui tend vers 0 quand $n \to +\infty$. Ainsi, la suite (f_n) converge simplement sur $\mathbb{R}$ vers la fonction nulle.

Pour tout $n \in \mathbb{N}^*$ et tout réel x, $f_n'(x) = n\dfrac{1-n^2x^2}{(1+n^2x^2)^2}$. f_n étant impaire, on en déduit facilement que $|f_n|$ atteint son maximum en $\dfrac{1}{n}$.

Donc $\|f_n\|_\infty = f_n\left(\dfrac{1}{n}\right) = \dfrac{1}{2}$, et il n'y a pas convergence uniforme sur $\mathbb{R}$.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ avec : $\forall\, x \in [0\,;1]$, $f_n(x) = nx^n \sin \pi x$.

Pour tout $n \in \mathbb{N}$ on a $f_n(1) = 0$; et pour $x \in [0\,;1[$ on a $|f_n(x)| \leqslant nx^n$ qui tend vers 0 quand $n \to +\infty$ (croissances comparées). La suite (f_n) converge donc simplement sur $[0\,;1]$ vers la fonction nulle.

Il paraît ici difficile de calculer explicitement le maximum de chaque $|f_n|$.

Cependant, pour tout $n \in \mathbb{N}^*$, on a $f_n\left(1-\dfrac{1}{n}\right) = n\left(1-\dfrac{1}{n}\right)^n \sin\left(\dfrac{\pi}{n}\right)$.

Puisque $\lim\limits_{n\to+\infty}\left(1-\dfrac{1}{n}\right)^n = \lim\limits_{n\to+\infty} \mathrm{e}^{n\ln\left(1-\frac{1}{n}\right)} = \dfrac{1}{\mathrm{e}}$ et que $\sin\left(\dfrac{\pi}{n}\right) \underset{n\to+\infty}{\sim} \dfrac{\pi}{n}$, on a : $\lim\limits_{n\to+\infty} f_n\left(1-\dfrac{1}{n}\right) = \dfrac{\pi}{\mathrm{e}}$.

Puisque $\|f_n\|_\infty^{[0;1]} \geqslant \left|f_n\left(1-\dfrac{1}{n}\right)\right|$, on ne peut pas avoir $\lim\limits_{n\to+\infty} \|f_n\|_\infty^{[0;1]} = 0$ donc la convergence de la suite (f_n) n'est pas uniforme sur $[0\,;1]$.

SF12.7 Raisonner par l'absurde

Lorsqu'il y a convergence uniforme d'une suite $(f_n)_{n\in\mathbb{N}}$ vers f, la limite f doit vérifier certaines propriétés parmi celles évoquées dans **[S12.2]**. En raisonnant par l'absurde, on peut donc ainsi montrer qu'il n'y a pas convergence uniforme.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies par $f_n(x) = x^n$ sur $[0\,;1]$.

De façon immédiate, la suite (f_n) converge simplement sur $[0\,;1]$ vers la fonction $f : x \mapsto \begin{cases} 0 & \text{si } x \in [0\,;1[\\ 1 & \text{si } x = 1\,. \end{cases}$

Les fonctions f_n étant continues sur $[0\,;1]$ alors que f ne l'est pas, il ne peut y avoir convergence uniforme sur $[0\,;1]$.

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies sur $[0\,;1]$ par :

$$f_n(x) = \frac{n}{nx+1} \text{ si } x \in]0\,;1] \quad \text{et} \quad f_n(0) = 0\,.$$

Il est facile de montrer que cette suite de fonctions converge simplement sur $[0\,;1]$ *vers la fonction* $f\colon x \mapsto \begin{cases} 0 & \text{si } x = 0 \\ \dfrac{1}{x} & \text{si } x \in \,]0\,;1]\,. \end{cases}$

Les f_n *n'étant pas continues, l'argument de l'exemple précédent ne peut s'appliquer.*

On remarque alors que pour tout $n \in \mathbb{N}$, $\lim\limits_{x\to 0^+} f_n(x) = n$ *existe. Si la convergence de la suite* (f_n) *vers* f *étant uniforme sur* $[0\,;1]$, *la suite* (n) *devrait avoir une limite dans* $\mathbb{R}$ *en vertu du théorème de la double limite, ce qui n'est pas le cas.*

On peut également remarquer que les f_n *sont bornées sur* $[0\,;1]$ *alors que* f *ne l'est pas : cela aussi contredit la convergence uniforme.*

Exemple

Étudier la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définies sur $[0\,;1]$ par : $f_n(x) = \dfrac{2^n x}{1 + n2^n x^2}$.

Nous vous laissons le soin de vérifier que cette suite de fonctions converge simplement sur $[0\,;1]$ *vers la fonction nulle (distinguer les cas* $x = 0$ *et* $x \in \,]0\,;1]$*).*

On calcule alors :

$$\int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \frac{2^n x}{1 + n2^n x^2}\,\mathrm{d}x = \frac{1}{2n}\left[\ln(1 + n2^n x^2\right]_0^1 = \frac{\ln(1 + n2^n)}{2n}\,.$$

Or :

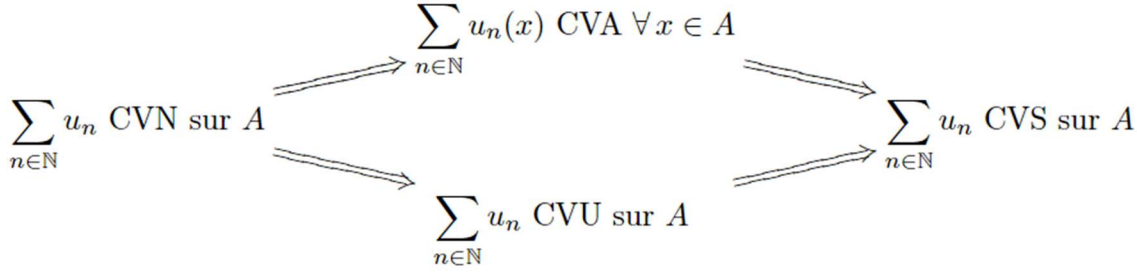
$$\frac{\ln(1 + n2^n)}{2n} \underset{n\to+\infty}{\sim} \frac{\ln(n2^n)}{2n} = \frac{\ln n + n\ln 2}{2n} \underset{n\to+\infty}{\sim} \frac{n\ln 2}{2n}\,,$$

donc $\lim\limits_{n\to+\infty} \int_0^1 f_n = \dfrac{\ln 2}{2}$, *alors que* $\int_0^1 \lim\limits_{n\to+\infty} f_n = 0$.

La suite (f_n) *ne peut donc pas converger uniformément vers la fonction nulle, en vertu de* **[S12.2]** *(interversion limite-intégrale).*

■ Comment étudier la convergence d'une série de fonctions ?

Pour étudier une série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$, il faut d'abord se rappeler les suites d'implications ci-dessous (aucune des réciproques n'étant vraie) ; en utilisant les abréviations CVS, CVA, CVU et CVN pour convergence simple, absolue, uniforme et normale respectivement, on a :



On pourra donc adopter le plan d'étude suivant (il s'agit là d'un plan général, qu'il faut bien sûr adapter à chaque cas et aux questions posées).

1. Chercher le domaine I de convergence simple.
2. Si la série converge normalement sur I, c'est-à-dire si $\sum_{n\in\mathbb{N}} \|u_n\|_\infty^I$ converge, alors il y a aussi convergence uniforme et absolue sur I.
3. Sinon, on peut étudier s'il y a convergence normale sur des parties convenables de I ; si oui, il y a convergence uniforme locale de la série, et on pourra appliquer les théorèmes sur ces parties.
4. Sinon, il ne reste plus qu'à étudier directement la convergence uniforme, c'est-à-dire étudier $\lim_{n\to\infty} \|R_n\|_\infty$.

✓ Parfois, la convergence normale de la série de fonctions sur le domaine donné est presque immédiate. Dans ce cas bien sûr, aucune étude supplémentaire n'est nécessaire.

✓ S'il y a convergence uniforme, puisque $u_n = R_{n-1} - R_n$, on aura $\|u_n\|_\infty \leqslant \|R_{n-1}\|_\infty + \|R_n\|_\infty$ donc nécessairement $\lim_{n\to+\infty} \|u_n\|_\infty = 0$.
On peut utiliser cette remarque pour montrer la *non* convergence uniforme de la série : si $\|u_n\|_\infty$ ne tend pas vers zéro, il ne peut pas y avoir convergence uniforme.

SF12.8 Étudier la convergence simple

Conformément à la définition **[S12.7]**, étudier la convergence simple d'une série de fonctions $\sum u_n$ revient à étudier la série $\sum u_n(x)$ pour x dans un ensemble convenable. On dispose donc de tous les outils habituels pour étudier une série : équivalents, développements limités, règles de comparaison, théorème sur les séries alternées etc.
Étudier la convergence simple équivaut à trouver le domaine de définition de la fonction $f = \sum_{n=0}^{+\infty} u_n$. Il s'agit de l'étude à faire en premier, sauf bien sûr si l'on peut prouver directement la convergence normale sur l'intervalle considéré (puisqu'elle implique tout le reste).

Exemple

Étudier la série de fonctions $\sum_{n\in\mathbb{N}^*} (-1)^n \frac{\mathrm{e}^{-nx}}{n+x}$.

Notons $u_n(x) = (-1)^n \frac{\mathrm{e}^{-nx}}{n+x}$.

- *On remarque déjà que, de toutes façons, cette série ne peut être définie lorsque* $x \in \mathbb{Z}_-^*$ *puisqu'alors il existe un entier* $n \in \mathbb{N}^*$ *tel que* $x + n = 0$.
- *Lorsque* x *est strictement négatif et n'appartient pas à* $\mathbb{Z}_-^*$, $\lim_{n\to+\infty} \frac{\mathrm{e}^{-nx}}{n+x} = +\infty$ *donc la série diverge grossièrement.*
- *Lorsque* $x = 0$, *il s'agit de la série harmonique alternée, dont on sait qu'elle est convergente.*

- *Enfin, si* $x > 0$, $\lim\limits_{n\to+\infty} n^2 u_n(x) = 0$ *d'après les résultats sur les croissances comparées, donc* $u_n(x) \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$. *Les théorèmes de comparaison sur les séries numériques permettent alors d'affirmer que la série est absolument convergente.*

Conclusion : la série converge pour tout $x \in \mathbb{R}_+$.

SF12.9 Étudier la convergence normale

Étudier la convergence normale d'une série de fonctions $\sum u_n$ sur un ensemble I, c'est étudier la convergence de la série numérique (à termes positifs) $\sum \|u_n\|_\infty^I$.

Pour cela, on peut soit calculer explicitement $\|u_n\|_\infty^I$, par exemple à l'aide d'une étude de fonctions, soit, et c'est en général plus simple, se contenter de majorer $|u_n(x)|$ pour $x \in I$ par un réel α_n *indépendant de* x tel que la série $\sum \alpha_n$ converge. En effet dans ce cas l'on aura $\|u_n\|_\infty^I \leqslant \alpha_n$ et les critères de comparaison sur les séries à termes réels positifs impliquent la convergence de la série $\sum \|u_n\|_\infty^I$.

✓ Bien sûr, si les u_n ne sont pas bornées sur I, au moins à partir d'un certain rang, il ne peut pas y avoir convergence normale !

Exemple

Étudier la série de fonctions $\sum\limits_{n\in\mathbb{N}^*} u_n$ où $u_n : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{x^\alpha}{(n+x)^\beta} \end{array} \right.$,

α et β étant des réels strictement positifs fixés.

- *Déterminons d'abord les valeurs des paramètres* α *et* β *pour lesquels il y a convergence simple.*
  - *Pour* $x = 0$, $u_n(0) = 0$ *pour tout* n *donc la série* $\sum u_n(0)$ *converge !*
  - *Pour* $x > 0$, $u_n(x) \underset{n\to+\infty}{\sim} \frac{x^\alpha}{n^\beta}$, *donc la série converge si et seulement si* $\beta > 1$.

  Il y a donc convergence simple sur $\mathbb{R}_+$ *si et seulement si* $\beta > 1$, *ce que l'on supposera désormais.*
- *Étudions maintenant la convergence normale sur* $\mathbb{R}_+$.
  - *Si* $\alpha > \beta$, $\lim\limits_{x\to+\infty} u_n(x) = +\infty$ *puisque, à* n *fixé,* $u_n(x) \underset{x\to+\infty}{\sim} \frac{x^\alpha}{x^\beta}$. *Les* u_n *n'étant pas bornées, il ne peut y avoir convergence normale sur* $\mathbb{R}_+$.
  - *Si* $\alpha = \beta$, $\lim\limits_{x\to+\infty} u_n(x) = 1$ *donc* $\|u_n\|_\infty^{\mathbb{R}_+} \geqslant 1$ *; la série* $\sum \|u_n\|_\infty^{\mathbb{R}_+}$ *est grossièrement divergente, et il n'y a pas convergence normale sur* $\mathbb{R}_+$.
  - *Supposons donc maintenant* $\alpha < \beta$. *Le calcul de* u_n', *que nous vous laissons le soin de faire, permet d'obtenir le tableau de variations suivant :*

x	0		$\frac{n\alpha}{\beta-\alpha}$		$+\infty$
$u_n'(x)$		$+$	0	$-$	
$u_n(x)$	0	↗		↘	0

On a donc :

$$\|u_n\|_\infty^{\mathbb{R}_+} = u_n\left(\frac{n\alpha}{\beta-\alpha}\right) = \frac{\left(\frac{n\alpha}{\beta-\alpha}\right)^\alpha}{\left(n+\frac{n\alpha}{\beta-\alpha}\right)^\beta} = \frac{\alpha^\alpha(\beta-\alpha)^{\beta-\alpha}}{\beta^\beta}\frac{1}{n^{\beta-\alpha}}\cdot$$

Par comparaison à une série de Riemann, on en déduit que la série $\sum \|u_n\|_\infty^{\mathbb{R}_+}$ *converge si et seulement si* $\beta-\alpha>1$.

- *Conclusion : la série de fonctions* $\sum_{n\in\mathbb{N}^*} u_n$ *converge simplement sur* $\mathbb{R}_+$ *si et seulement si* $\beta>1$ *et il y a convergence normale si et seulement si* $\beta>1+\alpha$.

SF12.10 Étudier la convergence uniforme

Pour étudier la convergence uniforme lorqu'il n'y a pas convergence normale, la seule solution consiste à considérer le reste R_n de la série (après avoir bien sûr étudié la convergence simple, sinon le reste n'est pas défini !) et à étudier si $\lim\limits_{n\to+\infty}\|R_n\|_\infty = 0$. Il faut donc tenter de majorer $|R_n(x)|$ indépendamment de x.

Pour cela, on peut essayer :

- d'utiliser la majoration du reste fournie par le critère spécial sur les séries alternées, si bien sûr celui-ci s'applique ;
- d'utiliser une méthode de comparaison avec une intégrale lorsque, pour x fixé, la suite $n\mapsto u_n(x)$ peut s'étendre en une fonction $t\mapsto u_t(x)$ intégrable ;
- de comparer à des séries dont on sait calculer la somme ;
- de majorer par la somme d'une série convergente, ou minorer par celle d'une série divergente.

Exemple

Étudier la série de fonctions $\sum_{n\in\mathbb{N}^*} u_n$ où $u_n(x)=\dfrac{1}{n+(n(x-n))^2}$ pour $x\in\mathbb{R}$.

- *Il est clair que* $\sup\limits_{x\in\mathbb{R}}|u_n(x)| = u_n(n) = \dfrac{1}{n}$ *; puisque la série* $\sum\dfrac{1}{n}$ *diverge, la série de fonctions* $\sum_{n\in\mathbb{N}^*} u_n$ *n'est pas normalement convergente sur* $\mathbb{R}$.

  ✓ *On pourrait cependant démontrer que cette série est normalement convergente sur tout intervalle de la forme* $]-\infty\,;A]$ *avec* $A\in\mathbb{R}$.

- *Démontrons que la série est uniformément convergente sur* $\mathbb{R}$.
  - *Il faut d'abord démontrer la convergence simple. Celle-ci découle immédiatement du fait que, à* x *fixé, on a* $u_n(x)\underset{n\to+\infty}{\sim}\dfrac{1}{n^4}$.
  - *Posons alors, pour tout* x *réel,* $R_n(x)=\sum\limits_{k=n+1}^{+\infty}u_k(x)$. *Dans cette somme, il y a au plus deux termes pour lesquels* $|x-k|$ *est* <1, *ceux correspondant éventuellement à* $k=\lfloor x\rfloor$ *ou* $k=\lfloor x+1\rfloor$ *; pour ces deux termes on aura* $u_k(x)\leqslant\dfrac{1}{k}\leqslant\dfrac{1}{n+1}$. *Pour tous les autres on aura* $|x-k|\geqslant 1$ *et alors* $u_k(x)=\dfrac{1}{k+(k(x-k))^2}\leqslant\dfrac{1}{k+k^2}\leqslant\dfrac{1}{k^2}$.

En regroupant les deux cas, on obtient l'inégalité :

$$\forall x \in \mathbb{R},\ 0 \leqslant R_n(x) \leqslant \frac{2}{n+1} + \sum_{k=n+1}^{+\infty} \frac{1}{k^2},$$

d'où : $\|R_n\|_\infty \leqslant \dfrac{2}{n+1} + \displaystyle\sum_{k=n+1}^{+\infty} \frac{1}{k^2}$.

Le terme de droite étant le reste d'une série convergente tend vers 0 quand $n \to +\infty$, *donc* $\lim\limits_{n\to+\infty} \|R_n\|_\infty = 0$, *ce qui prouve bien la convergence uniforme sur* $\mathbb{R}$.

Exemple

Étude de la série de fonctions $\displaystyle\sum_{n\geqslant 1} (-1)^{n-1} \ln\left(1+\frac{x}{n}\right)$ avec $x \geqslant 0$.

Posons, pour $n \geqslant 1$ *et* $x \geqslant 0$, $u_n(x) = (-1)^{n-1} \ln\left(1+\dfrac{x}{n}\right)$.

- *On note d'abord que les fonctions* u_n *sont bien définies sur* $\mathbb{R}_+$.
- *Il est facile de vérifier que, pour* $x \geqslant 0$ *fixé, la suite* $n \mapsto u_n(x)$ *vérifie les hypothèses du critère spécial sur les séries alternées. Cela prouve la convergence simple de la série de fonctions.*
- *Il ne peut y avoir convergence normale sur aucun intervalle inclus dans* $\mathbb{R}_+$, *puisque, sauf pour* $x = 0$, *la série n'est pas absolument convergente (*$|u_n(x)| \underset{n\to+\infty}{\sim} \dfrac{x}{n}$, *qui est le terme général d'une série divergente).*
- *Il ne peut pas y avoir non plus convergence uniforme sur* $\mathbb{R}_+$, *les* u_n *n'étant pas bornées.*
- *Cependant, le critère spécial sur les séries alternées donne aussi, en notant* R_n *le reste d'ordre* n *de la série :*

$$\forall x \geqslant 0,\ \forall n \in \mathbb{N}^*,\ |R_n(x)| \leqslant |u_{n+1}(x)| = \ln\left(1+\frac{x}{n+1}\right),$$

donc, si A *est un réel strictement positif on a :*

$$\|R_n\|_\infty^{[0;A]} \leqslant \ln\left(1+\frac{A}{n+1}\right) \leqslant \frac{A}{n+1}.$$

Ainsi, $\lim\limits_{n\to+\infty} \|R_n\|_\infty^{[0;A]} = 0$, *et il y a convergence uniforme sur* $[0\,;A]$.

Exemple

Étude de la série de fonctions $\displaystyle\sum_{n\in\mathbb{N}} \frac{x}{1+n^2x^2}$ $(x \geqslant 0)$.

Posons, pour $n \in \mathbb{N}$ *et* $x \in \mathbb{R}_+$, $u_n(x) = \dfrac{x}{1+n^2x^2}$.

- *La série converge simplement sur* $\mathbb{R}_+$ *; en effet, pour* $x = 0$, $u_n(0) = 0$ *donc la série converge, et pour* $x \neq 0$, $u_n(x) \underset{n\to+\infty}{\sim} \dfrac{1}{n^2x}$, *donc la série converge par comparaison à une série de Riemann.*

- *Pour $n \in \mathbb{N}^*$ on a $u_n\left(\frac{1}{n}\right) = \frac{1}{2n}$ donc $\|u_n\|_\infty^{\mathbb{R}_+} \geqslant \frac{1}{2n}$; puisque la série de terme général $\frac{1}{2n}$ diverge, la série de fonctions $\sum u_n$ n'est pas normalement convergente sur $\mathbb{R}_+$.*
  Il est cependant facile de montrer la convergence normale sur tout intervalle de la forme $[a\,;+\infty[$ avec $a > 0$ puisque :
  $$\forall\, n \in \mathbb{N}^*,\ \forall\, x \geqslant a,\ 0 \leqslant u_n(x) \leqslant \frac{x}{n^2x^2} = \frac{1}{n^2x} \leqslant \frac{1}{n^2a}\,,$$
  donc $\sum_{n\in\mathbb{N}} \|u_n\|_\infty^{[a;+\infty[}$ converge, par comparaison à une série de Riemann.
- *Montrons enfin qu'il n'y a pas convergence uniforme sur $\mathbb{R}_+$. Notons R_n le reste d'ordre n : $\forall\, x \geqslant 0$, $R_n(x) = \sum_{k=n+1}^{+\infty} \frac{x}{1+k^2x^2}$.*
  Pour $x \geqslant 0$ fixé et $t \in \mathbb{R}_+$, posons $f_x(t) = \frac{x}{1+t^2x^2}$. f_x est continue et décroissante sur $\mathbb{R}_+$ donc :
  $$\forall\, k \in \mathbb{N},\ u_k(x) = f_x(k) \geqslant \int_k^{k+1} f_x(t)\,\mathrm{d}t\,.$$
  f_x étant intégrable sur $\mathbb{R}_+$, on peut sommer ces inégalités pour k variant de $n+1$ à $+\infty$ et on obtient pour tout $x \geqslant 0$:
  $$R_n(x) \geqslant \int_{n+1}^{+\infty} \frac{x}{1+t^2x^2}\,\mathrm{d}t = \int_{(n+1)x}^{+\infty} \frac{\mathrm{d}u}{1+u^2} = \frac{\pi}{2} - \mathrm{Arctan}\big((n+1)x\big).$$
  En particulier, $R_n\left(\frac{1}{n+1}\right) \geqslant \frac{\pi}{4}$, donc $\|R_n\|_\infty^{\mathbb{R}_+}$ ne peut pas tendre vers 0 quand $n \to +\infty$, et il n'y a pas convergence uniforme sur $\mathbb{R}_+$.

■ Comment étudier la limite en un point de la somme d'une série de fonctions ?

SF12.11 Utiliser le théorème de la double limite

Il s'agit du cas le plus simple. Il faut seulement vérifier la convergence uniforme de la série au voisinage du point considéré.

Exemple

Étudier la convergence simple de la série de fonctions $\sum_{n\in\mathbb{N}} u_n$ où $u_n : x \mapsto \mathrm{e}^{-x\sqrt{n}}$.

On note S sa somme. Déterminer $\lim\limits_{x\to+\infty} S(x)$.

- *Les fonctions u_n sont définies sur $\mathbb{R}$.*
- *Si $x \leqslant 0$, $u_n(x)$ ne tend pas vers 0 lorsque $n \to +\infty$ donc la série $\sum_{n\in\mathbb{N}} u_n(x)$ diverge grossièrement.*
  Si $x > 0$, $\lim\limits_{n\to+\infty} n^2u_n(x) = 0$ donc $u_n(x) \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$. Les théorèmes de comparaison sur les séries à termes positifs permettent de conclure que la série converge simplement sur $\mathbb{R}_+^*$.

- *Il ne peut y avoir convergence normale sur $\mathbb{R}_+^*$ puisque $\|u_n\|_\infty^{\mathbb{R}_+^*} = 1$; cependant, pour la seconde question, seul importe ce qui se passe au voisinage de $+\infty$. Or, si a est un réel strictement positif fixé, on a $u_n(x) \leqslant \mathrm{e}^{-a\sqrt{n}}$ pour tout $x \geqslant a$, et puisque la série de terme général $\mathrm{e}^{-a\sqrt{n}}$ converge, la série $\sum_{n\in\mathbb{N}} u_n$ converge normalement sur $[a\,;+\infty[$.*

- *Pour tout $n \in \mathbb{N}$, $\lim\limits_{x\to+\infty} u_n(x) = \ell_n$ existe, et $\ell_n = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{sinon .} \end{cases}$*

  ✓ *Attention au cas $n = 0$ dans le calcul de cette limite ! C'est une erreur malheureusement fréquente. . .*

  Puisque la série de fonctions converge normalement donc uniformément sur $[a\,;+\infty[$, le théorème de la double limite s'applique et donne :

$$\lim_{x\to+\infty} S(x) = \sum_{n=0}^{+\infty} \ell_n = 1\,.$$

SF12.12 Utiliser des majorations et/ou minorations

Lorsque le théorème de la double limite ne s'applique pas, on peut essayer d'encadrer la somme de la série pour trouver la limite, ou de la minorer convenablement pour montrer la divergence.

Exemple

Pour tout $x > 0$ on pose $S(x) = \sum\limits_{n=0}^{+\infty} \dfrac{x}{1+n^2x^2}$. Déterminer $\lim\limits_{x\to0^+} S(x)$.

Cette série de fonctions a déjà été étudiée dans un exemple page 490 ; nous avions montré la convergence simple mais non uniforme de la série sur $\mathbb{R}_+$.

La convergence n'étant pas uniforme, le théorème de la double limite ne peut s'appliquer. Nous allons donc encadrer S à l'aide d'une comparaison avec une intégrale.

Pour $x > 0$ fixé et $t \in \mathbb{R}_+$, posons $f_x(t) = \dfrac{x}{1+t^2x^2}$. f_x est continue et décroissante sur $\mathbb{R}_+$ donc :

$$\forall\, k \in \mathbb{N},\ u_k(x) = f_x(k) \geqslant \int_k^{k+1} f_x(t)\,\mathrm{d}t\,.$$

f_x étant intégrable sur $\mathbb{R}_+$, on peut sommer ces inégalités pour k variant de 0 à $+\infty$ et on obtient pour tout $x > 0$:

$$S(x) \geqslant \int_0^{+\infty} \frac{x}{1+t^2x^2}\,\mathrm{d}t = \int_0^{+\infty} \frac{\mathrm{d}u}{1+u^2} = \frac{\pi}{2}\,.$$

On a également :

$$\forall\, k \in \mathbb{N}^*,\ u_k(x) = f_x(k) \leqslant \int_{k-1}^{k} f_x(t)\,\mathrm{d}t\,,$$

d'où en sommant pour k variant de 1 à $+\infty$, compte tenu de $u_0(x) = x$:

$$S(x) - x \leqslant \int_0^{+\infty} \frac{x}{1+t^2x^2}\,\mathrm{d}t = \frac{\pi}{2}\,,$$

et l'on obtient finalement l'encadrement :

$$\forall x > 0,\ \frac{\pi}{2} \leqslant S(x) \leqslant \frac{\pi}{2} + x\,.$$

On en déduit $\lim\limits_{x\to 0^+} S(x) = \dfrac{\pi}{2}$ *(et, puisque* $S(0) = 0$*, cela prouve que* S *n'est pas continue en* 0*).*

Exemple

Montrer que la fonction $f : x \mapsto \sum\limits_{n=0}^{+\infty} \ln(1+x^n)$ est définie sur $[0\,;1[$, et déterminer sa limite en 1^-.

- *L'étude du domaine de définition de* f*, c'est-à-dire de la convergence simple de la série de fonctions, ne doit pas poser de problèmes.*
  - *Pour* $x = 0$*, la série converge trivialement et* $f(0) = 0$*.*
  - *Pour* $x \in]0\,;1[$*,* $\lim\limits_{n\to+\infty} x^n = 0$ *donc* $\ln(1+x^n) \underset{n\to+\infty}{\sim} x^n$*, et la série converge par comparaison à une série géométrique.*
- *Puisque l'on devine aisément que* $\lim\limits_{x\to 1^-} f(x) = +\infty$*, nous allons chercher à minorer* f*.*
  Pour tout $N \in \mathbb{N}$ *et tout* $x \in [0\,;1[$ *on a*

$$\sum_{n=0}^{N} \ln(1+x^n) = \ln\left(\prod_{n=0}^{N}(1+x^n)\right) \geqslant \ln(1+x+\cdots+x^N)$$

puisque, en développant, le produit $\prod\limits_{n=0}^{N}(1+x^n)$ *comporte l'expression* $1+x+\cdots+x^N$ *plus des termes positifs.*

En utilisant la formule donnant la somme des termes d'une suite géométrique, on obtient :

$$\forall x \in [0\,;1[,\ \sum_{n=0}^{N} \ln(1+x^n) \geqslant \ln\left(\frac{1-x^{N+1}}{1-x}\right),$$

puis en faisant tendre N *vers* $+\infty$ *:* $\forall x \in [0\,;1[,\ f(x) \geqslant \ln\left(\dfrac{1}{1-x}\right)$.

On en déduit alors immédiatement $\lim\limits_{x\to 1^-} f(x) = +\infty$.

SF12.13 Utiliser le théorème de la limite monotone

Ce théorème, vu en première année, a été rappelé en **[S9.1]** page 102.

Exemple

Étudier la convergence simple de la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$ où $u_n : x \mapsto \mathrm{e}^{-x\sqrt{n}}$.

On note S sa somme. Déterminer $\lim\limits_{x\to 0^+} S(x)$.

Cette série de fonctions a déjà été étudiée dans un exemple page 491, où nous avions montré sa convergence simple sur $\mathbb{R}_+^*$.

Puisque pour tout entier n on a $u_n(0) = 1$, on peut légitimement penser que S va tendre vers $+\infty$ lorsque $x \to 0^+$. Démontrons-le rigoureusement.

Les fonctions $x \mapsto u_n(x) = \mathrm{e}^{-x\sqrt{n}}$ étant décroissantes sur $\mathbb{R}_+$, il en est de même de S ; en effet, si $0 < x \leqslant y$ on a $u_n(x) \geqslant u_n(y)$ pour tout n d'où par sommation des inégalités $S(x) \geqslant S(y)$.

> ✓ *Vous remarquerez sur cet exemple qu'il n'est pas toujours utile de calculer la dérivée pour étudier le sens de variations d'une fonction !*
> *Cela aurait été ici une grosse perte de temps puisqu'il fallait appliquer le théorème de dérivation d'une série de fonctions, donc en vérifier soigneusement toutes les hypothèses.*

La fonction S étant décroissante sur $\mathbb{R}_+^*$, le théorème de la limite monotone indique que $\lim\limits_{x\to 0^+} S(x)$ existe si et seulement si S est majorée. Montrons que ce n'est pas le cas en raisonnant par l'absurde.
Supposons S majorée sur $\mathbb{R}_+^*$; dans ce cas, il existe un réel M tel que :

$$\forall\, x > 0,\ S(x) = \sum_{n=0}^{+\infty} \mathrm{e}^{-x\sqrt{n}} \leqslant M\,,$$

ce qui implique :

$$\forall\, x > 0,\ \forall\, N \in \mathbb{N},\ \sum_{n=0}^{N} \mathrm{e}^{-x\sqrt{n}} \leqslant M\,.$$

Dans cette dernière inégalité, qui met en jeu une somme finie, *on peut faire tendre x vers 0^+ ce qui donne :*

$$\forall\, N \in \mathbb{N},\ \sum_{n=0}^{N} 1 \leqslant M\,,$$

ou encore : $\forall\, N \in \mathbb{N},\ N + 1 \leqslant M$ avec M fixé, ce qui est absurde.
Conclusion : S étant décroissante non majorée, on a : $\lim\limits_{x\to 0^+} S(x) = +\infty$.

SF12.14 Revenir à la définition de la limite

Cette méthode est surtout efficace dans le cas d'une limite infinie. Exposons-là sur un exemple.

Exemple

Pour tout $x > 1$, on pose $\zeta(x) = \sum\limits_{n=1}^{+\infty} \dfrac{1}{n^x}$. Montrer que $\lim\limits_{x\to 1^+} \zeta(x) = +\infty$.

Notons déjà que, d'après le cours sur les séries de Riemann, la fonction ζ est bien définie sur $]1\,;+\infty[$.

Soit A un réel positif quelconque. Puisque la série $\sum \dfrac{1}{n}$ est une série à termes positifs divergente, par définition de la limite infinie il existe un entier N tel que $\sum\limits_{n=1}^{N} \dfrac{1}{n} \geqslant A + 1$ $(*)$.

D'autre part, $\lim\limits_{x\to 1^+} \sum\limits_{n=1}^{N} \dfrac{1}{n^x} = \sum\limits_{n=1}^{N} \dfrac{1}{n}$ *donc par définition de la limite il existe un réel* $\alpha > 0$ *tel que :*

$$\forall\, x \in]1\,;\alpha[,\ \left|\sum_{n=1}^{N} \frac{1}{n^x} - \sum_{n=1}^{N} \frac{1}{n}\right| < 1 \ (**)$$

et en combinant $(*)$ *et* $(**)$ *on obtient :* $\forall\, x \in]1\,;\alpha[,\ \zeta(x) \geqslant \sum\limits_{n=1}^{N} \dfrac{1}{n^x} \geqslant A$,
ce qui signifie, par définition de la limite, que $\lim\limits_{x\to 1^+} \zeta(x) = +\infty$.

■ Comment étudier la continuité de la somme d'une série de fonctions ?

SF12.15 Utiliser le théorème du cours

Puisque la continuité est une notion locale, pour démontrer que la somme d'une série de fonctions continues est continue en un point a, il suffit de démontrer la convergence uniforme de la série au voisinage de a. Il s'agit de la méthode employée dans pratiquement tous les cas.
Nous ne donnons pas ici d'exemples particuliers puisque cette méthode a déjà été mise en œuvre dans de nombreux exemples précédents, et le sera encore dans ceux qui vont suivre.

SF12.16 Revenir à la définition

Bien sûr, on ne reviendra à la définition que si la méthode précédente ne s'applique pas, c'est-à-dire en cas de non convergence uniforme locale.
La définition de la continuité d'une fonction f en un point a est :

$$\lim_{x\to a} f(x) = f(a)\,.$$

Pour étudier la continuité, on peut donc se ramener à un calcul de limite. Nous ne détaillons pas davantage : c'est ce que nous avons fait dans l'exemple traité en **[SF12.12]** page 492.

■ Comment étudier la dérivabilité de la somme d'une série de fonctions?

SF12.17 Utiliser le théorème du cours

Ce théorème indique que, si une série de fonctions de classe $\mathscr{C}^1$ converge simplement, la fonction somme sera de classe $\mathscr{C}^1$ sous réserve que la série des dérivées converge uniformément, au moins localement.

Exemple

On pose, pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}^+$: $u_n(x) = \dfrac{\mathrm{e}^{-nx}}{(n+x)^2}$.

Montrer que $S\colon x \mapsto \displaystyle\sum_{n=1}^{+\infty} u_n(x)$ est continue sur $\mathbb{R}_+$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$.

- *On remarque d'abord que, puisque $n \geqslant 1$, les fonctions u_n sont bien toutes définies sur $\mathbb{R}_+$.*
  Pour tout $x \geqslant 0$, on a $0 \leqslant u_n(x) \leqslant \dfrac{1}{n^2}$ donc $\|u_n\|_\infty \leqslant \dfrac{1}{n^2}$. Par comparaison à une série de Riemann, la série $\sum u_n$ est normalement donc uniformément convergente sur $\mathbb{R}_+$. Les u_n étant continues, il en résulte que S est définie et continue sur $\mathbb{R}_+$.

  ✓ *Notez que nous avons ici directement démontré la convergence normale, ce qui rend inutile l'étude de la convergence simple.*

- *Les u_n sont de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$, mais on n'a plus convergence normale de la série des dérivées ; en effet, $u_n'(x) = \dfrac{-n\,\mathrm{e}^{-nx}}{(n+x)^2} - \dfrac{2\,\mathrm{e}^{-nx}}{(n+x)^3}$ d'où $|u_n'(x)| = \dfrac{n\,\mathrm{e}^{-nx}}{(n+x)^2} + \dfrac{2\,\mathrm{e}^{-nx}}{(n+x)^3}$, donc chaque fonction $x \mapsto |u_n'(x)|$ est décroissante (comme somme et produit de fonctions décroissantes positives).*
  Par conséquent, $\|u_n'\|_\infty^{\mathbb{R}_+} = |u_n'(0)| = \dfrac{1}{n} + \dfrac{2}{n^3} \geqslant \dfrac{1}{n}$, et la série $\sum \|u_n'\|_\infty^{\mathbb{R}_+}$ diverge.
  En revanche, si a est un réel strictement positif, sur $[a\,;+\infty[$ on a :

$$\|u_n'\|_\infty^{[a;+\infty[[} = |u_n'(a)| \underset{n\to+\infty}{\sim} \frac{\mathrm{e}^{-na}}{n},$$

  donc $\|u_n'\|_\infty^{[a;+\infty[[} = o\left(\dfrac{1}{n^2}\right)$ et la série $\sum \|u_n'\|_\infty^{[a;+\infty[}$ converge.
  Il y a donc convergence normale et par suite convergence uniforme sur $[a\,;+\infty[$. Ainsi, S est de classe $\mathscr{C}^1$ sur tout intervalle $[a\,;+\infty[$ donc est de classe $\mathscr{C}^1$ sur la réunion de ces intervalles c'est-à-dire $\mathbb{R}_+^*$.

SF12.18 Revenir à la définition

On peut revenir à la définition de la dérivée (limite du taux d'accroissement) lorsque la méthode précédente ne s'applique pas, ou pour montrer que la somme de la série n'est pas dérivable.

Exemple

Pour tout entier $n \geqslant 1$ et tout $x \geqslant 0$, on pose $u_n(x) = \dfrac{x}{n(1+nx)^2}$.

Montrer que la fonction $S\colon x \mapsto \displaystyle\sum_{n=1}^{+\infty} u_n(x)$ est continue sur $\mathbb{R}_+$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$. Est-elle dérivable en 0 ?

- *Les u_n sont bien toutes définies sur $\mathbb{R}_+$.*

- *Pour $x = 0$ la série converge.*
  Pour $x > 0$, $u_n(x) \underset{n\to+\infty}{\sim} \dfrac{1}{n^3 x}$; or la série $\sum \dfrac{1}{n^3}$ est à termes positifs et convergente (série de Riemann) donc la série $\sum u_n(x)$ converge par comparaison de séries à termes positifs.
  Il y a donc convergence simple sur $\mathbb{R}_+$.
- *Les u_n sont de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$ et pour $x \geqslant 0$, $u_n'(x) = \dfrac{1-nx}{n(1+nx)^3}$.*
  u_n' s'annule en $x = \dfrac{1}{n}$ et l'étude des variations de u_n montre que $\|u_n\|_\infty^{\mathbb{R}_+} = u_n\left(\dfrac{1}{n}\right) = \dfrac{1}{4n^2}$.
  On en déduit la convergence normale de la série sur $\mathbb{R}_+$. Les u_n étant continues sur $\mathbb{R}_+$, cela permet de conclure quant à la continuité de S.
- *Soit $[a\,;b]$ un segment de $\mathbb{R}_+^*$. Alors on a :*
$$\forall\, x \in [a\,;b],\ |u_n'(x)| \leqslant \frac{nb+1}{n(1+na)^3} = O\left(\frac{1}{n^3}\right),$$
  ce qui montre la convergence normale donc uniforme de $\sum u_n'$ sur $[a\,;b]$. S est donc de classe $\mathscr{C}^1$ sur tout segment $[a\,;b]$ de $\mathbb{R}_+^*$ donc sur $\mathbb{R}_+^*$.
- *Montrons que S n'est pas dérivable en 0. Pour cela on étudie la limite du rapport $\dfrac{S(x)-S(0)}{x-0} = \dfrac{S(x)}{x}$ quand $x \to 0^+$.*
  Pour $x > 0$, $\dfrac{S(x)}{x} = \displaystyle\sum_{n=1}^{+\infty} \frac{1}{n(1+nx)^2}$. Pour montrer que $\lim\limits_{x\to0^+} \dfrac{S(x)}{x} = +\infty$, on peut utiliser le théorème de la limite monotone comme en **[SF12.13]**. *Exposons une autre méthode possible.*
  Pour $n \in \mathbb{N}^*$ et $0 < x \leqslant \dfrac{1}{n}$ on a $1 + nx \leqslant 2$ donc $\dfrac{1}{n(1+nx)^2} \geqslant \dfrac{1}{4n}$. On aura donc, pour tout $N \in \mathbb{N}^*$:
$$\forall\, x \in \left]0\,;\frac{1}{N}\right[,\ \frac{S(x)}{x} \geqslant \sum_{n=1}^{N} \frac{1}{4n}.$$
  En revenant à la définition de la limite, comme dans l'exemple traité en **[SF12.14]** *page 494, on en conclut que $\lim\limits_{x\to0^+} \dfrac{S(x)}{x} = +\infty$. Donc S n'est pas dérivable en 0.*

SF12.19 Utiliser le théorème de prolongement de la dérivée

Ce théorème a été rappelé en **[S9.1]** page 103.

Exemple

Pour tout entier $n \in \mathbb{N}$ et tout $x \geqslant 0$, on pose $u_n(x) = (-1)^n \dfrac{\mathrm{e}^{-nx}}{n+1}$. Montrer que la fonction $S : x \mapsto \displaystyle\sum_{n=0}^{+\infty} u_n(x)$ est continue sur $\mathbb{R}_+$ et de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$. Est-elle dérivable en 0 ?

- *La série converge simplement sur $\mathbb{R}_+$ car pour tout $x \geqslant 0$ la suite $n \mapsto u_n(x)$ vérifie les hypothèses du critère spécial des séries alternées.*

- *Ce même critère implique, en notant R_n le reste d'ordre n :*

$$\forall x \geqslant 0,\ |R_n(x)| \leqslant \frac{\mathrm{e}^{-(n+1)x}}{n+2} \leqslant \frac{1}{n+2}\,,$$

donc $\lim\limits_{n\to+\infty} \|R_n\|_\infty^{\mathbb{R}_+} = 0$: la série converge uniformément sur $\mathbb{R}_+$, ce qui implique que S y est continue.

- *Les u_n sont de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$ et $u_n'(x) = (-1)^{n+1}\dfrac{n}{n+1}\mathrm{e}^{-nx}$. Pour tout $a > 0$ et tout $x \in [a\,;+\infty[$ on a $|u_n'(x)| \leqslant \mathrm{e}^{-na}$, ce qui implique la convergence normale (donc uniforme) de la série sur $[a\,;+\infty[$.*
On peut donc en conclure que S est de classe $\mathscr{C}^1$ sur $[a\,;+\infty[$ pour tout $a > 0$ donc sur $\mathbb{R}_+^*$, et que :

$$\forall x > 0,\ S'(x) = \sum_{n=0}^{+\infty} (-1)^{n+1}\frac{n}{n+1}\mathrm{e}^{-nx}\,.$$

- *Dérivabilité en 0.*
Puisque la série $\sum u_n'(0)$ diverge, la série $\sum u_n'$ n'est pas uniformément convergente sur $\mathbb{R}_+$ (car sinon le théorème de la double limite pourrait s'appliquer).
Cependant, nous allons montrer que S est dérivable en 0.
On avait, pour tout $x > 0$: $S'(x) = \sum\limits_{n=0}^{+\infty} (-1)^{n+1}\dfrac{n}{n+1}\mathrm{e}^{-nx}$. En utilisant l'astuce classique consistant à écrire $n = (n+1) - 1$ on aura donc :

$$\begin{aligned}\forall x > 0,\ S'(x) &= \sum_{n=0}^{+\infty} (-1)^{n+1}\frac{(n+1)-1}{n+1}\mathrm{e}^{-nx}\\ &= \sum_{n=0}^{+\infty} (-1)^{n+1}\mathrm{e}^{-nx} - \sum_{n=0}^{+\infty} (-1)^{n+1}\frac{\mathrm{e}^{-nx}}{n+1}\end{aligned}$$

(on peut en effet « couper la somme en 2 » puisque les séries écrites sont toutes convergentes pour $x > 0$). La première de ces sommes est une série géométrique de raison $-\mathrm{e}^{-x}$, la seconde série est celle de S donc :

$$\forall x > 0,\ S'(x) = -\frac{1}{1+\mathrm{e}^{-x}} + S(x)\,.$$

S étant continue sur $\mathbb{R}_+$, on en déduit $\lim\limits_{x\to 0^+} S'(x) = -\dfrac{1}{2} + S(0)$. Le théorème de prolongement de la dérivée permet alors d'en déduire que S est dérivable en 0 et que S est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$.

■ Comment intégrer terme à terme une série de fonctions ?

SF12.20 Utiliser le théorème du cours

Ce théorème permet d'intervertir $\sum$ et $\int$ lorsque la série de fonctions converge uniformément sur un *segment*.

Exemple

Étudier la série de fonctions $\sum_{n\geqslant 1} \frac{a^n \cos nx}{n}$ où a est un réel appartenant à $]-1\,;1[$.

En déduire la valeur de l'intégrale $I = \int_0^\pi \ln(a^2 - 2a\cos x + 1)\,\mathrm{d}x$.

On notera, pour tout $x \in \mathbb{R}$ et tout $n \geqslant 1$: $u_n(x) = \frac{a^n \cos nx}{n}$.

- Domaine de définition et continuité

  Pour tout $x \in \mathbb{R}$, on a $|u_n(x)| \leqslant \frac{|a|^n}{n} \leqslant |a|^n$, donc $\|u_n\|_\infty^{\mathbb{R}} \leqslant |a|^n$. Par comparaison à une série géométrique, on en déduit que la série de fonctions $\sum_{n\geqslant 1} u_n$ est normalement donc uniformément convergente sur $\mathbb{R}$.

  On peut donc poser, pour tout $x \in \mathbb{R}$: $S(x) = \sum_{n=1}^{+\infty} \frac{a^n \cos nx}{n}$, et les u_n étant continues sur $\mathbb{R}$, il en est de même de S par convergence uniforme.

- Dérivabilité

  Pour tout $n \in \mathbb{N}^*$, la fonction u_n est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et, pour tout $x \in \mathbb{R}$, $u_n'(x) = -a^n \sin nx$. On a donc $\|u_n'\|_\infty^{\mathbb{R}} = |a|^n$, et la série de fonctions $\sum_{n\geqslant 1} u_n'$ converge normalement donc uniformément sur $\mathbb{R}$.

  Le théorème de dérivation d'une série de fonctions permet alors d'affirmer que S est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et que :

$$\forall\, x \in \mathbb{R},\ S'(x) = \sum_{n=1}^{+\infty} -a^n \sin nx = \sum_{n=0}^{+\infty} -a^n \sin nx\,.$$

  Calculons cette somme :

$$\begin{aligned} S'(x) &= -\mathcal{I}m\left(\sum_{n=0}^{+\infty} a^n \mathrm{e}^{\mathrm{i}nx}\right) = -\mathcal{I}m\left(\sum_{n=0}^{+\infty} (a\mathrm{e}^{\mathrm{i}x})^n\right) \\ &= -\mathcal{I}m\left(\frac{1}{1-a\mathrm{e}^{\mathrm{i}x}}\right) = -\frac{a\sin x}{a^2 - 2a\cos x + 1}\,. \end{aligned}$$

  Or, pour $x = 0$, $S(0) = \sum_{n=1}^{+\infty} \frac{a^n}{n} = -\ln(1-a)$ (développement en série entière de la fonction $\ln$ sur $]-1\,;1[$), donc on a, pour tout $x \in \mathbb{R}$:

$$\begin{aligned} S(x) &= S(0) + \int_0^x S'(t)\,\mathrm{d}t = -\ln(1-a) + \int_0^x -\frac{a\sin t}{a^2 - 2a\cos t + 1}\,\mathrm{d}t \\ &= -\ln(1-a) - \frac{1}{2}\left[\ln\left|a^2 - 2a\cos t + 1\right|\right]_0^x \\ &= -\ln(1-a) + \frac{1}{2}\ln((1-a)^2) - \frac{1}{2}\ln\left|a^2 - 2a\cos x + 1\right| \\ &= -\frac{1}{2}\ln(a^2 - 2a\cos x + 1)\,. \end{aligned}$$

- *D'après le calcul précédent,* $I = \int_0^\pi \ln(a^2 - 2a\cos x + 1)\,\mathrm{d}x = -2\int_0^\pi S(x)\,\mathrm{d}x$.

La série $\displaystyle\sum_{n=1}^{+\infty} \frac{a^n \cos nx}{n}$ *étant normalement donc uniformément convergente sur* $\mathbb{R}$*, donc sur* $[0,\pi]$*, on peut intégrer terme à terme :*

$$\int_0^\pi S(x)\,\mathrm{d}x = \sum_{n=1}^{+\infty} \int_0^\pi \frac{a^n \cos nx}{n}\,\mathrm{d}x = 0$$

d'où : $I = 0$.

SF12.21 Montrer que l'intégrale du reste tend vers zéro

La méthode que nous allons exposer ci-dessous peut permettre de conclure lorsqu'aucun théorème du cours ne s'applique (ou que vous n'avez pas trouvé comment les appliquer !).
Pour montrer que :

$$\int_I \left(\sum_{n=0}^{+\infty} u_n(x)\right)\mathrm{d}x = \sum_{n=0}^{+\infty} \int_I u_n(x)\,\mathrm{d}x \quad (1)$$

(en supposant que tout ce qui est écrit a un sens, c'est-à-dire après avoir vérifié l'existence des séries et des intégrales), il suffit de montrer que :

$$\lim_{n\to+\infty} \int_I \left(\sum_{k=n+1}^{+\infty} u_k(x)\right)\mathrm{d}x = 0 \quad (2)$$

En effet, par linéarité de l'intégrale on a :

$$\begin{aligned}\int_I \left(\sum_{n=0}^{+\infty} u_n(x)\right)\mathrm{d}x &= \int_I \left(\sum_{k=0}^{n} u_k(x)\right)\mathrm{d}x + \int_I \left(\sum_{k=n+1}^{+\infty} u_k(x)\right)\mathrm{d}x \\ &= \sum_{k=0}^{n} \int_I u_k(x)\,\mathrm{d}x + \int_I \left(\sum_{k=n+1}^{+\infty} u_k(x)\right)\mathrm{d}x\,,\end{aligned}$$

donc la propriété (2) implique la convergence de la série $\displaystyle\sum_{n\geqslant 0} \int_I u_n$ et prouve aussi l'égalité (1).

Exemple

Soit $(a_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles strictement positives, croissante et telle que $\displaystyle\lim_{n\to+\infty} a_n = +\infty$.

Démontrer que $\displaystyle\int_0^1 \left(\sum_{n=0}^{+\infty} (-1)^n x^{a_n}\right)\mathrm{d}x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{a_n+1}$.

✓ Attention : cet énoncé est assez abrupt ! Il faudra bien sûr commencer par vérifier l'existence de tout ce qui y est écrit.

On notera, pour tout $x \in [0\,;1]$ *et tout* $n \in \mathbb{N}$, $u_n(x) = (-1)^n x^{a_n}$ *et* $S_n = \displaystyle\sum_{k=0}^{n} u_k$.

- *Nous vous laissons le soin de vérifier que, pour* $x \in [0\,;1[$*, la suite* $n \mapsto u_n(x)$ *vérifie les hypothèses du critère spécial sur les séries alternées,*

donc la série $\sum u_n$ converge simplement sur $[0\,;1[$ (mais elle ne converge pas en 1). On pourra donc poser, pour $0 \leqslant x < 1$: $S(x) = \sum_{n=0}^{+\infty} u_n(x)$ et pour tout $n \in \mathbb{N}$, $R_n(x) = \sum_{k=n+1}^{+\infty} u_k(x)$.

- *Toujours à l'aide du critère spécial sur les séries alternées, on a :*
$$\forall\, x \in [0\,;1[,\ |R_n(x)| \leqslant |u_{n+1}(x)| = x^{a_{n+1}}\,,$$
donc si b est un réel compris strictement entre 0 et 1 on a :
$$\|R_n\|_\infty^{[0;b]} = b^{a_{n+1}}\,,\ \text{d'où}\ \lim_{n\to+\infty} \|R_n\|_\infty^{[0;b]} = 0\,,$$
et la série est normalement convergente sur tout segment $[0\,;b] \subset [0\,;1[$. Cela implique la continuité des applications S et R_n sur $[0\,;1[$.
- *La majoration précédente $|R_n(x)| \leqslant x^{a_{n+1}}$ implique, d'après les règles de comparaison pour les intégrales de fonctions positives, que R_n est intégrable sur $[0\,;1[$ et l'on a :*
$$\left|\int_0^1 R_n(x)\,\mathrm{d}x\right| \leqslant \int_0^1 |R_n(x)|\,\mathrm{d}x \leqslant \int_0^1 x^{a_{n+1}}\,\mathrm{d}x = \frac{1}{a_{n+1}+1}\,,$$
donc $\lim_{n\to+\infty} \int_0^1 R_n = 0$ (1).
- *Pour tout $n \in \mathbb{N}$, S_n est continue sur $[0\,;1]$ donc y est intégrable et*
$$\int_0^1 S_n = \sum_{k=0}^{n} \int_0^1 u_k = \sum_{k=0}^{n} \frac{(-1)^k}{a_k+1}$$
puis $S = S_n + R_n$ est intégrable sur $[0\,;1[$ et
$$\int_0^1 S = \int_0^1 S_n + \int_0^1 R_n = \sum_{k=0}^{n} \frac{(-1)^k}{a_k+1} + \int_0^1 R_n \quad (2)$$
Puisque la série de terme général $\frac{(-1)^k}{a_k+1}$ est convergente d'après le critère spécial sur les séries alternées, on déduit de (1) et (2) que $\int_0^1 S = \sum_{k=0}^{+\infty} \frac{(-1)^k}{a_k+1}$.

SF12.22 Utiliser d'autres théorèmes du cours

Dans le chapitre relatif à l'intégration, nous verrons deux autres théorèmes qui permettent d'intégrer terme à terme une série de fonctions. Il s'agit des théorèmes de convergence dominée **[S14.9]** et d'intégration terme à terme **[S14.10]**.
Ces théorèmes sont les plus utilisés car il permettent de conclure lorsqu'il y a seulement convergence simple ou lorsque l'intervalle d'intégration n'est pas un segment, et nous verrons en détail leur utilisation lors de l'étude du thème 14.

■ Comment trouver un équivalent de la somme d'une série de fonctions?

SF12.23 Utiliser une comparaison avec une intégrale

Il s'agit d'une des méthodes les plus utilisées pour obtenir un encadrement de la somme d'une série (que ce soit une série de fonctions ou une série numérique). Si cet encadrement est assez précis, il peut permettre d'obtenir un équivalent.

Exemple

Pour $x \in \mathbb{R}$ on pose $f(x) = \sum_{n=1}^{+\infty} \frac{1}{n^2+x^2}$. Déterminer la limite et un équivalent de $f(x)$ lorsque x tend vers $+\infty$.

Notons, pour tout $x \in \mathbb{R}$ et tout $n \geqslant 1$: $u_n(x) = \frac{1}{x^2+n^2}$.

- *Pour tout x réel, on a $|u_n(x)| \leqslant \frac{1}{n^2}$ donc $\|u_n\|_\infty^{\mathbb{R}} \leqslant \frac{1}{n^2}$ et d'après les théorèmes de comparaison sur les séries à termes positifs, la série $\sum \|u_n\|_\infty$ est convergente.*
  Ainsi la série de fonctions est-elle normalement donc uniformément convergente sur $\mathbb{R}$. Le théorème de la double limite donne alors $\lim_{x\to+\infty} f(x) = \sum_{n=1}^{+\infty} \lim_{x\to+\infty} u_n(x) = 0$.
- *Pour tout $x > 0$, la fonction $t \mapsto \frac{1}{x^2+t^2}$ est continue décroissante et positive sur $\mathbb{R}_+$. La méthode de comparaison série-intégrale donne alors :*
$$\forall n \geqslant 1\ ,\ \int_n^{n+1} \frac{\mathrm{d}t}{x^2+t^2} \leqslant \frac{1}{x^2+n^2} \leqslant \int_{n-1}^{n} \frac{\mathrm{d}t}{x^2+t^2}\ ,$$
donc en sommant, puisque les intégrales convergent :
$$\int_1^{+\infty} \frac{\mathrm{d}t}{x^2+t^2} \leqslant S(x) \leqslant \int_0^{+\infty} \frac{\mathrm{d}t}{x^2+t^2}\ .$$
Une primitive de $t \mapsto \frac{1}{x^2+t^2}$ étant $t \mapsto \frac{1}{x}\operatorname{Arctan}\frac{t}{x}$, on en tire :
$$\forall x > 0\ ,\ \frac{\pi}{2x} - \frac{1}{x}\operatorname{Arctan}\frac{1}{x} \leqslant S(x) \leqslant \frac{\pi}{2x}\ ,$$
d'où l'on déduit facilement : $S(x) \underset{x\to+\infty}{\sim} \frac{\pi}{2x}$ (puisque, d'après le théorème des gendarmes, le quotient de ces deux fonctions tend vers 1).

SF12.24 Utiliser le théorème de la double limite

Exemple

On pose $S(x) = \sum_{n\geqslant 1} (-1)^{n-1} \ln\left(1+\frac{x}{n}\right)$ pour $x \geqslant 0$.

Déterminer la limite et un équivalent de S en 0^+.

- *Il y a convergence simple de la série de fonctions sur $\mathbb{R}_+$ d'après le critère spécial sur les séries alternées. La somme S de la série est donc bien définie sur $\mathbb{R}_+$.*
- *Si l'on note $R_n(x)$ le reste d'ordre n de la série, on a :*
$$|R_n(x)| \leqslant |u_{n+1}(x)| = \ln\left(1+\frac{x}{n+1}\right) \leqslant \frac{x}{n+1},$$
donc si a est un réel positif $\|R_n\|_\infty^{[0;a]} \leqslant \dfrac{a}{n+1} \underset{n\to+\infty}{\longrightarrow} 0$, ce qui prouve que la série converge uniformément sur tout segment de la forme $[0\,;a]$. Les u_n étant continues, il en résulte que S est continue sur $\mathbb{R}_+$.
- *S étant continue sur $\mathbb{R}_+$, on a $\lim\limits_{x\to0^+} S(x) = S(0) = 0$. De plus, la série $\sum\limits_{n\geqslant1} \dfrac{u_n(x)}{x}$ converge uniformément sur $\mathbb{R}_+^*$ puisque pour tout $x>0$, le reste d'ordre n de cette série est majoré en valeur absolue par $\dfrac{1}{n+1}$ d'après le calcul précédent.*
D'après le théorème d'interversion des limites on peut donc écrire :
$$\lim_{x\to0^+}\frac{S(x)}{x} = \sum_{n=1}^{+\infty}\lim_{x\to0^+}\frac{u_n(x)}{x} = \sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n} = \ln 2,$$
ce qui prouve que $S(x) \underset{x\to0^+}{\sim} x\ln 2$.

> ✓ *Une autre solution possible était d'utiliser un développement limité d'ordre 1, après l'avoir justifié bien sûr.*
> *Un exemple de cette méthode sera donné en* **[SF12.27]**.

SF12.25 Remarquer un terme prépondérant

Si dans la somme d'une série de fonctions on remarque un terme visiblement prépondérant sur tous les autres, on peut essayer de démontrer que la somme est équivalente à ce terme.

Exemple

Pour $x\in\mathbb{R}$ on pose $f(x) = \sum\limits_{n=1}^{+\infty} \mathrm{e}^{-n^2x}$. Déterminer la limite et un équivalent de $f(x)$ quand $x\to+\infty$.

- *Déjà, il est facile de démontrer que cette série de fonctions converge simplement sur $\mathbb{R}^*$ puisque, par exemple, on a $\mathrm{e}^{-n^2x} \underset{n\to+\infty}{=} o\left(\dfrac{1}{n^2}\right)$ pour $x\neq0$. De plus, la série diverge grossièrement pour $x=0$ donc f est définie sur $\mathbb{R}^*$.*
- *La série converge normalement sur tout intervalle de la forme $[a\,;+\infty[$ avec $a>0$ puisque, pour tout $x\geqslant a$ on a $\mathrm{e}^{-n^2x} \leqslant \mathrm{e}^{-n^2a}$ qui est le terme général d'une série convergente. Il y a donc aussi convergence uniforme sur cet intervalle, et le théorème de la double limite permet de conclure :*
$$\lim_{x\to+\infty} f(x) = \sum_{n=1}^{+\infty}\lim_{x\to+\infty}\mathrm{e}^{-n^2x} = 0.$$

- *Puisque* $f(x) = \mathrm{e}^{-x} + \mathrm{e}^{-4x} + \mathrm{e}^{-9x} + \cdots$, *on peut légitimement penser que* $f(x)$ *sera équivalent à* e^{-x} *quand* $x \to +\infty$. *Démontons-le rigoureusement.*

  Pour cela, on calcule, pour tout $x > 0$: $\dfrac{f(x)}{\mathrm{e}^{-x}} = 1 + \displaystyle\sum_{n=2}^{+\infty} \mathrm{e}^{(1-n^2)x}$. *Pour les mêmes raisons que ci-dessus, la série de fonctions* $\displaystyle\sum_{n\geqslant 2} \mathrm{e}^{(1-n^2)x}$ *converge normalement donc uniformément sur tout intervalle de la forme* $[a\,;+\infty[$ *donc là encore le théorème de la double limite s'applique :*

$$\lim_{x\to+\infty} \left(\sum_{n=2}^{+\infty} \mathrm{e}^{(1-n^2)x} \right) = 0\,.$$

  Il en résulte $\lim\limits_{x\to+\infty} \dfrac{f(x)}{\mathrm{e}^{-x}} = 1$ *donc* $f(x) \underset{x\to+\infty}{\sim} \mathrm{e}^{-x}$.

SF12.26 Sommer des équivalents

Lors du cours de première année, vous avez souvent du entendre la phrase : « On ne peut pas additionner des équivalents ! »
Cette affirmation péremptoire est surtout faite pour éviter de faire des erreurs grossières lorsque l'on débute dans la manipulation des équivalents. En fait, il est souvent possible de faire cette opération, à condition de la justifier rigoureusement en revenant à la définition.

Exemple

Soit $f : [0\,;1[\longrightarrow \mathbb{R}$ définie par $f(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{nx+1}$.

Donner un équivalent de $f(x)$ lorsque x tend vers 1^-.

Déjà, la définition a bien un sens puisque, pour $x \in [0\,;1[$, $0 \leqslant \dfrac{x^n}{nx+1} \leqslant x^n$ *qui est le terme général d'une série géométrique convergente.*

Quand x *tend vers* 1, *on a* $\dfrac{x^n}{nx+1} \sim \dfrac{x^n}{n+1}$ *; cela incite à comparer* $f(x)$ *à* $g(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{x^n}{n+1}$ *(qui est une série entière dont on sait calculer la somme). Or :*

$$\forall\, x \in [0\,;1[,\ 0 \leqslant \frac{x^n}{nx+1} - \frac{x^n}{n+1} = \frac{n}{n+1}\frac{x^n(1-x)}{nx+1} \leqslant \frac{x^n}{nx+1}(1-x)\,,$$

et en sommant ces inégalités pour n *variant de* 0 *à* $+\infty$:

$$\forall\, x \in [0\,;1[,\ 0 \leqslant f(x) - g(x) \leqslant (1-x)f(x)\,.$$

Donc quand x *tend vers* 1^-, *on a* $f(x) - g(x) = o(f(x))$ *soit* $f(x) \underset{x\to1^-}{\sim} g(x)$.

Pour terminer, on remarque que $xg(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1} = -\ln(1-x)$ *(développement en série entière qui doit être connu), donc pour* $x \in]0\,;1[$, $g(x) = -\dfrac{\ln(1-x)}{x}$ *et enfin* $f(x) \underset{x\to1^-}{\sim} -\ln(1-x)$.

■ Comment obtenir un développement asymptotique de la somme d'une série de fonctions ?

Cette problématique généralise la précédente.

SF12.27 Utiliser la formule de Taylor-Young

Il s'agit de la façon la plus simple d'obtenir un développement limité au voisinage d'un point, lorsque la fonction est suffisamment dérivable en ce point.

Exemple

Étudier la série de fonctions $\sum_{n\geqslant 1} \frac{1}{x^2+n^2}$.

Donner un développement limité de la somme S à l'ordre 2 au voisinage de 0.

Notons, pour tout $x \in \mathbb{R}$ et tout $n \geqslant 1$: $u_n(x) = \frac{1}{x^2+n^2}$.

- *Nous avons déjà montré, dans un exemple traité page 502, que cette série est normalement donc uniformément convergente sur $\mathbb{R}$.*
- *Les u_n sont de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et pour tout x, $u_n'(x) = \frac{-2x}{(x^2+n^2)^2}$. Sur tout intervalle $[-A\,;A]$ avec $A>0$ on a $\|u_n'\|_\infty^{[-A;A]} \leqslant \frac{2A}{n^4}$ donc la série $\sum \|u_n'\|_\infty^{[-A;A]}$ est convergente.*
  Il en résulte que la série de fonctions $\sum u_n'$ est normalement donc uniformément convergente sur tout segment inclus dans $\mathbb{R}$. D'après le théorème de dérivation, on en déduit que la fonction S est de classe $\mathscr{C}^1$ sur $\mathbb{R}$ et que pour tout x réel : $S'(x) = \sum_{n=1}^{+\infty} u_n'(x)$.
- *En appliquant de nouveau ce même théorème de dérivation à la série de fonctions $\sum u_n'$ on montre facilement que S est de classe $\mathscr{C}^2$ sur $\mathbb{R}$ et que pour tout x réel : $S''(x) = \sum_{n=1}^{+\infty} u_n''(x)$.*
  Le développement limité demandé s'obtient alors directement par la formule de Taylor-Young :

$$S(x) = S(0) + xS'(0) + \frac{x^2}{2}S''(0) + o\big(x^2\big) = \frac{\pi^2}{6} - \frac{\pi^4}{90}x^2 + o\big(x^2\big)\,.$$

✓ *On a utilisé ici les relations :*

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad ; \quad \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}\,.$$

Ces formules ne figurent pas au programme officiel. Si vous ne vous rappelez pas les valeurs de ces sommes, vous pouvez quand même finir l'exercice, en les nommant ($\zeta(2)$ et $\zeta(4)$, par exemple...)

SF12.28 Sommer des développements limités

Il est bien sûr possible de sommer des développements limités, mais il faut alors majorer rigoureusement les restes.

Exemple

Pour tout $n \in \mathbb{N}^*$ et tout $x \geqslant 0$, on pose : $u_n(x) = \dfrac{(-1)^n}{nx+1}$.

Étudier la convergence de la série de fonctions $\sum_{n\geqslant 1} u_n$.

On note S sa somme. Montrer que $\lim\limits_{x\to+\infty} S(x) = 0$ et donner un développement asymptotique à deux termes de S en $+\infty$.

- *S définie sur $\mathbb{R}_+^*$: cela découle facilement du critère spécial sur les séries alternées.*
- *À l'aide de la majoration du reste, on démontre que la série converge uniformément sur tout intervalle $[a\,;+\infty[$; on peut donc appliquer le théorème de la double limite, qui donne le résultat demandé (nous n'avons pas détaillé car ce raisonnement a déjà été fait plusieurs fois).*
- *Pour $n \in \mathbb{N}^*$ fixé, on a : $\dfrac{1}{nx+1} = \dfrac{1}{nx}\left(1+\dfrac{1}{nx}\right)^{-1}$ et, puisque $\dfrac{1}{nx}$ tend vers 0 lorsque $x \to +\infty$, on peut songer à faire un développement limité, en utilisant celui de $\dfrac{1}{1+h}$ au voisinage de 0. Mais pour pouvoir ensuite faire la somme pour n variant de 1 à $+\infty$, il faut avoir une estimation assez précise du reste. Pour cela, on calcule :*

$$\forall\, h > 0,\ \varepsilon(h) = \frac{1}{1+h} - (1-h) = \frac{h^2}{1+h}\cdot$$

Ainsi $\dfrac{1}{1+h} = 1 - h + \varepsilon(h)$ avec $|\varepsilon(h)| \leqslant h^2$ (ce résultat est donc plus précis qu'un simple développement limité).

On peut alors écrire, pour tout $n \geqslant 1$ et tout $x > 0$:

$$u_n(x) = \frac{(-1)^n}{nx}\frac{1}{1+\frac{1}{nx}} = \frac{(-1)^n}{nx}\left(1 - \frac{1}{nx} + \varepsilon\left(\frac{1}{nx}\right)\right),$$

et il est possible de sommer ces égalités puisque toutes les séries convergent :

$$\forall\, x > 0,\ S(x) = \sum_{n=1}^{+\infty}\frac{(-1)^n}{nx} - \sum_{n=1}^{+\infty}\frac{(-1)^n}{n^2x^2} + \sum_{n=1}^{+\infty} v_n(x)$$

où $v_n(x) = \dfrac{(-1)^n}{nx}\varepsilon\left(\dfrac{1}{nx}\right)$ donc $|v_n(x)| \leqslant \dfrac{1}{n^3x^3}$ puis $\left|\sum\limits_{n=1}^{+\infty} v_n(x)\right| \leqslant \dfrac{K}{x^3}$ avec $K = \sum\limits_{n=1}^{+\infty}\dfrac{1}{n^3}$.

Compte tenu de $\sum\limits_{n=1}^{+\infty}\dfrac{(-1)^n}{n} = -\ln 2$ et $\sum\limits_{n=1}^{+\infty}\dfrac{(-1)^n}{n^2} = -\dfrac{\pi^2}{12}$, on obtient le développement asymptotique :

$$S(x) \underset{x\to+\infty}{=} -\frac{\ln 2}{x} + \frac{\pi^2}{12x^2} + O\left(\frac{1}{x^3}\right).$$

✓ *Le calcul de* $\displaystyle\sum_{n=1}^{+\infty}\frac{(-1)^n}{n^2}$ *est classique. Nous le détaillons ci-dessous.*

Notons $\displaystyle S=\sum_{n=1}^{+\infty}\frac{1}{n^2}\cdot$ *En séparant dans cette somme les termes d'indices pairs et impairs (toutes les séries convergent), on obtient :*

$$S=\sum_{k=1}^{+\infty}\frac{1}{(2k-1)^2}+\sum_{k=1}^{+\infty}\frac{1}{(2k)^2}=\sum_{k=1}^{+\infty}\frac{1}{(2k-1)^2}+\frac{1}{4}S\,.$$

Notons $\displaystyle S'=\sum_{n=1}^{+\infty}\frac{(-1)^n}{n^2}\cdot$ *De la même façon, on a :*

$$S'=-\sum_{k=1}^{+\infty}\frac{1}{(2k-1)^2}+\sum_{k=1}^{+\infty}\frac{1}{(2k)^2}=-\sum_{k=1}^{+\infty}\frac{1}{(2k-1)^2}+\frac{1}{4}S\,.$$

On en déduit $S+S'=\frac{1}{2}S$ *d'où* $S'=-\frac{1}{2}S=-\frac{\pi^2}{12}$ *puisque* $S=\frac{\pi^2}{6}\cdot$

À vous de jouer !

Exercice 12.1 **SF12.4**

Soit $(P_n)_{n\in\mathbb{N}}$ une suite de fonctions polynômes convergeant uniformément sur $\mathbb{R}$ vers une fonction f.

Montrer que f est une fonction polynôme.

Exercice 12.2 **SF12.1** **SF12.3** **SF12.4**

Soient $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ deux suites de fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$.

1) a) On suppose que (f_n) converge uniformément vers f sur $\mathbb{R}$, et que (g_n) converge simplement vers g.
Montrer que la suite $(f_n\circ g_n)$ converge simplement vers $f\circ g$.

b) Montrer que ce résultat peut tomber en défaut si l'on suppose seulement que (f_n) converge simplement vers f.

2) On suppose maintenant que la suite (f_n) converge uniformément vers f sur $\mathbb{R}$, que la suite (g_n) converge uniformément vers g sur $\mathbb{R}$ et que f est uniformément continue sur $\mathbb{R}$.

a) Montrer que la suite $(f_n\circ g_n)$ converge uniformément vers $f\circ g$ sur $\mathbb{R}$.

b) Montrer que ce résultat peut tomber en défaut si l'on ne suppose pas f uniformément continue.

Exercice 12.3 **SF12.4**

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions définies sur un intervalle $[a\,;b]$, à valeurs dans un espace vectoriel normé F, qui converge simplement vers une fonction f sur cet intervalle.
On suppose qu'il existe un réel $k \geqslant 0$ tel que, pour tout $n \in \mathbb{N}$, f_n est k-lipschitzienne.

Montrer que la suite (f_n) converge uniformément vers f sur $[a\,;b]$.

Indication : on commencera par traiter le cas où $f = 0$, et on introduira une subdivision bien choisie de l'intervalle $[a\,;b]$.

Exercice 12.4 **SF12.1** **SF12.2**

Étudier la suite de fonctions $(f_n)_{n\geqslant 1}$ définies sur $\mathbb{R}_+^*$ par :

$$\forall\, x > 0\ ,\ f_n(x) = \frac{\sin(nx)}{n\sqrt{x}}\,.$$

Exercice 12.5 **SF12.1** **SF12.3**

Étudier la suite de fonctions $(f_n)_{n\geqslant 1}$ définies sur $\mathbb{R}_+$ par :

$$f_n(x) = \begin{cases} \left(1-\dfrac{x}{n}\right)^n & \text{si } \ 0 \leqslant x < n\,, \\ 0 & \text{si } x \geqslant n\,. \end{cases}$$

Exercice 12.6 **SF12.8** **SF12.9** **SF12.10**

Soit $(a_n)_{n\in\mathbb{N}^*}$ une suite réelle décroissante et à valeurs positives.
Pour tout $n \in \mathbb{N}^*$ et tout $x \in [0\,;1]$, on pose :

$$u_n(x) = a_n x^n (1-x)\,.$$

1) Montrer que la série $\sum u_n$ converge simplement sur $[0\,;1]$.

2) Montrer que la convergence est normale sur $[0\,;1]$ si et seulement si la série $\displaystyle\sum_{n\in\mathbb{N}^*} \frac{a_n}{n}$ converge.

3) Montrer que la convergence est uniforme sur $[0\,;1]$ si et seulement si la suite $(a_n)_{n\in\mathbb{N}^*}$ converge vers 0.

Exercice 12.7 **SF12.8** **SF12.10** **SF12.20**

On pose : $u_n(x) = (-1)^{n+1} x^{2n+2} \ln x$ pour $x \in\,]0\,;1]$ et $u_n(0) = 0\,.$

1) Calculer : $\displaystyle\sum_{n=0}^{+\infty} u_n(x)\,.$

2) Montrer que la série de fonctions $\sum_{n\in\mathbb{N}^*} u_n$ converge uniformément sur $[0\,;1]$.

3) En déduire l'égalité : $\displaystyle\int_0^1 \frac{\ln x}{1+x^2}\,\mathrm{d}x = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{(2n+1)^2}$.

Exercice 12.8 SF12.11 SF12.17 SF12.23 SF12.24 SF12.26

1) Quel est le domaine de définition de $S : x \mapsto \displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2x+n}$?

2) Montrer que S est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$.

3) Déterminer un équivalent de S en $+\infty$.

4) Déterminer un équivalent de S en 0^+.

Exercice 12.9 SF12.8 SF12.25 SF12.26

Soit f définie sur $\mathbb{R}_+^*$ par $f(x) = \displaystyle\sum_{n=1}^{+\infty} \frac{1}{\operatorname{sh} nx}$.

Déterminer un équivalent de $f(x)$ quand x tend vers 0^+ et quand x tend vers $+\infty$.

Exercice 12.10 SF12.8 SF12.10 SF12.11 SF12.26

Pour tout $n \in \mathbb{N}^*$ et tout $x \geqslant 0$ on pose $u_n(x) = \dfrac{(-1)^n}{\sqrt{nx+1}}$.

1) Étudier la convergence de la série de fonctions $\sum_{n\geqslant 1} u_n$.

2) On note S sa somme. Montrer que $\lim_{x\to+\infty} S(x) = 0$ et donner un développement asymptotique à un terme de S en $+\infty$.

Thème 13 - Séries entières

■ Comment déterminer le rayon de convergence d'une série entière ?

SF13.1 Revenir à la définition

La définition du rayon de convergence d'une série entière à partir du lemme d'Abel ou les définitions équivalentes qui s'ensuivent (cf. **[S13.1]**) permettent d'obtenir directement certains résultats concernant le rayon de convergence R d'une série entière $\sum_{n\geqslant 0} a_n z^n$.

- Si la série $\sum_{n\geqslant 0} a_n z_0^n$ converge pour un certain z_0, alors $R \geqslant |z_0|$.
- S'il existe z_0 tel que $\lim_{n\to+\infty} a_n z_0^n = 0$, alors $R \geqslant |z_0|$.
- S'il existe z_0 tel que la suite $(a_n z_0^n)$ soit bornée, alors $R \geqslant |z_0|$.
- Si la série $\sum_{n\geqslant 0} a_n z_1^n$ diverge pour un certain z_1, alors $R \leqslant |z_1|$.
- S'il existe z_1 tel que $\lim_{n\to+\infty} a_n z_1^n \neq 0$, alors $R \leqslant |z_1|$.
- S'il existe z_1 tel que la suite $(a_n z_1^n)$ ne soit pas bornée, alors $R \leqslant |z_1|$.
- S'il existe z_2 tel que la série $\sum_{n\geqslant 0} a_n z_2^n$ soit convergente mais pas absolument convergente, alors $R = |z_2|$.
- S'il existe z_2 tel que la la suite $(a_n z_2^n)$ soit bornée mais ne converge pas vers 0, alors $R = |z_2|$.
- Si la suite (a_n) a une limite ℓ *non nulle*, alors $R = 1$.

> ✓ Bien sûr, il ne faut pas apprendre tous les cas précédents pas cœur ! Mais il faut savoir les retrouver immédiatement avec un peu de réflexion à partir des caractérisations du rayon de convergence rappelées en **[S13.1]**.

Exemple

Déterminer le rayon de convergence R de la série entière $\sum_{n\geqslant 2} \dfrac{x^n}{n\ln n}$.

- *Puisque* $\lim_{n\to+\infty} \dfrac{1}{n\ln n} = 0$, *on a* $R \geqslant 1$.
- *Puisque pour* $x = 1$ *la série* $\sum_{n\geqslant 2} \dfrac{1}{n\ln n}$ *est divergente (par comparaison avec l'intégrale divergente* $\int_2^{+\infty} \dfrac{\mathrm{d}t}{t\ln t}$*), on a* $R \leqslant 1$.

  On pouvait aussi remarquer que pour $x > 1$, $\lim_{n\to+\infty} \dfrac{x^n}{n\ln n} = +\infty$ *; la série est alors divergente, d'où* $R \leqslant 1$.

En conclusion : $R = 1$.

Exemple

Déterminer le rayon de convergence de la série entière $\sum_{n\geqslant 1} n^{-\lfloor\sqrt{n}\rfloor} z^n$.

✓ Nous avons constaté que beaucoup d'entre vous sont très gênés dès qu'intervient la fonction partie entière dans un exercice. Mais dans la plupart des cas, on arrive à surmonter la difficulté en utilisant la propriété

$$\forall\, x\in\mathbb{R},\ \lfloor x\rfloor \leqslant x < \lfloor x\rfloor + 1\,,$$

ou bien la propriété suivante, qui en découle : $\lfloor x\rfloor \underset{x\to+\infty}{\sim} x$.

Posons ici $u_n = a_n x^n$ pour $n\in\mathbb{N}^*$ et $x\in\mathbb{R}^*$.

On a $\ln|u_n| = -\lfloor\sqrt{n}\rfloor \ln n + n\ln|x|$. Or $\lfloor\sqrt{n}\rfloor \underset{n\to+\infty}{\sim} \sqrt{n}$, donc si $|x|\neq 1$, $\ln|u_n| \underset{n\to+\infty}{\sim} n\ln|x|$ car $\sqrt{n}\ln n$ est négligeable devant n par les croissances comparées. On aura donc :

$$\lim_{n\to+\infty} \ln|u_n| = \begin{cases} -\infty & si\ |x|<1 \\ +\infty & si\ |x|>1 \end{cases} \quad puis \quad \lim_{n\to+\infty} |u_n| = \begin{cases} 0 & si\ |x|<1 \\ +\infty & si\ |x|>1\,. \end{cases}$$

On en déduit que $R = 1$.

Exemple

Rayon de convergence de la série entière $\sum a_n z^n$, où $a_n = \displaystyle\int_0^1 \left(\frac{1+t^2}{2}\right)^n \mathrm{d}t$.

L'encadrement $t\leqslant \dfrac{1+t^2}{2} \leqslant 1$ pour $0\leqslant t\leqslant 1$ donne $\dfrac{1}{n+1}\leqslant a_n\leqslant 1$.

L'inégalité de gauche implique que la série $\sum a_n$ diverge (par comparaison à la série harmonique), donc $R\leqslant 1$.

L'inégalité de droite prouve que la suite (a_n) est bornée, donc $R\geqslant 1$.

En conclusion, $R = 1$.

Exemple

Déterminer le rayon de convergence de la série entière $\sum a_n z^n$ où a_n vérifie :

$$a_{3p} = \frac{1}{p^2+1}\,,\ a_{3p+1} = \frac{1}{p!}\,,\ a_{3p+2} = a^p\ (a\in\mathbb{R}_+^*)\,.$$

Pour tout $z\in\mathbb{C}$, la suite $(a_n z^n)$ est bornée si et seulement si les trois suites $(a_{3p}z^{3p})$, $(a_{3p+1}z^{3p+1})$ et $(a_{3p+2}z^{3p+2})$ le sont. Or :

- *$(a_{3p}z^{3p})$ est bornée si et seulement si $|z|\leqslant 1$;*
- *$(a_{3p+1}z^{3p+1})$ est bornée pour tout $z\in\mathbb{C}$;*
- *$a_{3p+2}z^{3p+2} = a^p z^{3p+2} = z^2\,(\sqrt[3]{a}z)^{3p}$, donc $(a_{3p+2}z^{3p+2})$ est bornée si et seulement si $|\sqrt[3]{a}z|\leqslant 1$ c'est-à-dire si $|z|\leqslant \dfrac{1}{\sqrt[3]{a}}$.*

On en déduit que $R = \min\left(1, \dfrac{1}{\sqrt[3]{a}}\right)$.

SF13.2 Utiliser la règle de d'Alembert

Le critère de d'Alembert pour les séries entières doit être utilisé *avec la plus extrême précaution*, pour deux raisons :

- il ne s'applique que lorsque TOUS les coefficients de la série entière sont non nuls ;
- il ne s'applique alors que SI le rapport $\left|\frac{a_{n+1}}{a_n}\right|$ admet une limite (éventuellement infinie) lorsque $n \to +\infty$.

Dans la grande majorité des cas, on utilisera plutôt la règle de d'Alembert pour les séries numériques. De plus, l'utilisation de cette règle ne doit pas être systématique. Elle est par contre très pratique lorsque le terme a_n contient des exponentielles, des puissances ou des factorielles.

Exemple

Déterminer le rayon de convergence de la série entière $\sum_{n\geqslant 1} \frac{n!}{n^n} z^n$.

En posant $a_n = \dfrac{n!}{n^n}$, *on a* $a_n \neq 0$ *pour tout* n *et :*

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)!}{n!}\frac{n^n}{(n+1)^{n+1}} = \left(\frac{n}{n+1}\right)^n = \frac{1}{\left(1+\frac{1}{n}\right)^n}\cdot$$

Or $\lim\limits_{n\to+\infty}\left(1+\dfrac{1}{n}\right)^n = \lim\limits_{n\to+\infty} \mathrm{e}^{n\ln\left(1+\frac{1}{n}\right)} = \mathrm{e}$, *donc* $\lim\limits_{n\to+\infty}\left|\dfrac{a_{n+1}}{a_n}\right| = \dfrac{1}{\mathrm{e}}\cdot$

Le rayon de convergence cherché est donc égal à e.

Exemple

Déterminer le rayon de convergence de la série entière $\sum_{n\in\mathbb{N}} n! z^{n^2}$.

En posant ici $u_n = n!z^{n^2}$, *on a pour* $z \neq 0$,

$$\left|\frac{u_{n+1}}{u_n}\right| = (n+1)\,|z|^{(n+1)^2-n^2} = (n+1)\,|z|^{2n+1}\ .$$

On a donc, d'après les croissances comparées des suites usuelles,

$$\lim_{n\to+\infty}\left|\frac{u_{n+1}}{u_n}\right| = \begin{cases} 0 & \text{si } |z|<1 \\ +\infty & \text{si } |z|\geqslant 1\,. \end{cases}$$

D'après la règle de d'Alembert pour les séries à termes positifs :

- *si* $|z| \geqslant 1$, *la série* $\sum |u_n|$ *diverge ;*
- *si* $|z| < 1$, *cette série converge.*

Le rayon de convergence cherché est donc égal à 1.

Exemple

Déterminer le rayon de convergence de la série entière $\sum_{n\in\mathbb{N}} \operatorname{ch} n z^{2n}$.

En posant $u_n = \operatorname{ch} n z^{2n}$, *on a, pour* $z \neq 0$,

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{\operatorname{ch}(n+1)}{\operatorname{ch} n}|z|^2 \underset{n\to+\infty}{\sim} \frac{\mathrm{e}^{n+1}}{\mathrm{e}^n}|z|^2$$

(en utilisant l'équivalent $\operatorname{ch} x \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^x}{2}$*) donc* $\lim\limits_{n\to+\infty}\left|\frac{u_{n+1}}{u_n}\right| = \mathrm{e}\,|z|^2$.

D'après la règle de d'Alembert pour les séries à termes positifs, $R = \frac{1}{\sqrt{\mathrm{e}}}$.

✓ Dans les deux exemples précédents, il y avait bien des coefficients a_n nuls pour certaines valeurs de n (séries entières dites *lacunaires*), mais cela n'empêche pas d'appliquer la règle de d'Alembert pour les séries numériques. Il était par contre impossible d'appliquer le critère de d'Alembert spécifique aux séries entières.

SF13.3 Utiliser la comparaison avec une autre série entière

Soient $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons de convergence respectifs R_a et R_b.
Si $|a_n| \geqslant |b_n|$ (au moins à partir d'un certain rang) ou si $a_n = O(b_n)$ (ou $a_n = o(b_n)$), alors $R_a \leqslant R_b$.

En encadrant judicieusement $|a_n|$, ou en utilisant simultanément l'une des méthodes vues dans le savoir-faire précédent, cette propriété permet parfois d'obtenir le rayon de convergence à l'aide d'une double inégalité.

Exemple

Déterminer le rayon de convergence R de la série entière $\sum \mathrm{e}^{\sin n} z^n$.

Pour tout n *on a* $\mathrm{e}^{-1} \leqslant \mathrm{e}^{\sin n} \leqslant \mathrm{e}$. *Donc* R *est compris entre les rayons de convergence des séries* $\sum \mathrm{e}^{-1} z^n$ *et* $\sum \mathrm{e} z^n$. *Ces deux séries ayant un rayon de convergence égal à* 1 *(séries géométriques de raison* z*), on en déduit* $R = 1$.

SF13.4 Utiliser des équivalents

Si $|a_n| \underset{n\to+\infty}{\sim} b_n$, alors les séries entières $\sum\limits_{n\geqslant 0} a_n z^n$ et $\sum\limits_{n\geqslant 0} b_n z^n$ ont même rayon de convergence.

Pour utiliser ce critère, on pourra donc chercher un équivalent « simple » (par exemple à l'aide d'un développement limité) de $|a_n|$, puis éventuellement appliquer à la série entière $\sum\limits_{n\geqslant 0} b_n z^n$ les méthodes précédentes.

Exemple

Déterminer le rayon de convergence R de la série entière $\sum \ln(n!) z^n$.

Notons $a_n = \ln(n!)$ *pour* $n \in \mathbb{N}^*$.

- $a_n = \sum\limits_{k=1}^{n} \ln k = \sum\limits_{k=2}^{n} \ln k$. *Nous allons chercher un équivalent de* a_n *en utilisant la méthode de comparaison à une intégrale.*

La fonction $t \mapsto \ln t$ étant croissante sur $\mathbb{R}_+^*$ on a :

$$\forall\, k \geqslant 2,\ \int_{k-1}^{k} \ln t\,\mathrm{d}t \leqslant \ln k \leqslant \int_{k}^{k+1} \ln t\,\mathrm{d}t\,,$$

donc en sommant :

$$\forall\, n \geqslant 2,\ \int_{1}^{n} \ln t\,\mathrm{d}t \leqslant \sum_{k=2}^{n} \ln k \leqslant \int_{2}^{n+1} \ln t\,\mathrm{d}t\,,$$

et puisque $\int \ln t\,\mathrm{d}t = t\ln t - t$ il est facile d'en déduire $a_n \underset{n\to+\infty}{\sim} n\ln n$.

✓ *On peut d'ailleurs obtenir plus rapidement cet équivalent en utilisant la formule de Stirling.*

- *On en déduit que le rayon de convergence R de la série entière $\sum a_n z^n$ est le même que celui de la série entière $\sum n\ln n z^n$. Or la suite $(n\ln n)$ n'est pas bornée donc $R \leqslant 1$ et si $|z|<1$, $\lim\limits_{n\to+\infty} n\ln n z^n = 0$ donc $R \geqslant 1$, et finalement : $R = 1$.*
- *Autre solution possible*

  Puisque $n! = 1\cdot 2\cdots n$ on a $n \leqslant n! \leqslant n^n$ donc en utilisant **[SF13.3]** *R est compris entre les rayons de convergence des deux séries entières $\sum \ln n z^n$ et $\sum n\ln n z^n$, qui sont tous deux égaux à 1, d'où le résultat.*

SF13.5 Utiliser les opérations sur les séries entières

On peut parfois utiliser les théorèmes du cours qui donnent le rayon de convergence d'une série dérivée ou d'une série primitive **[S13.5]**, ou qui donnent une inégalité sur le rayon de convergence d'une série somme ou produit de deux autres **[S13.2]**.

Exemple

Soit $\theta \in \mathbb{R}$. Déterminer le rayon de convergence R de la série entière $\sum \cos(n\theta)x^n$ puis celui R' de la série entière $\displaystyle\sum_{n\geqslant 1} \frac{\cos(n\theta)}{n} x^n$.

- *On a pour tout n $|\cos(n\theta)| \leqslant 1$ donc $R \geqslant 1$ (en utilisant* **[SF13.1]** *ou* **[SF13.3]***).*
- *La suite $\big(\cos(n\theta)\big)_{n\in\mathbb{N}}$ ne converge pas vers 0 : en effet, si tel était le cas, on aurait aussi $\lim\limits_{n\to+\infty} \cos(2n\theta) = 0$ (suite extraite) ; or $\cos(2n\theta) = 2\cos^2(n\theta) - 1$ doit aussi tendre vers -1, d'où une contradiction.*

  La série $\sum \cos(n\theta)$ est donc grossièrement divergente, ce qui montre que $R \leqslant 1$ et finalement $R = 1$.
- *La série entière $\displaystyle\sum_{n\geqslant 1} \frac{\cos(n\theta)}{n} x^n$ a même rayon de convergence que sa série dérivée, donc $R' = R$.*

Exemple

Soit $\sum a_n z^n$ une série entière de rayon de convergence R. Que peut-on dire du rayon de convergence R' de la série entière $\sum b_n z^n$ avec $b_n = \sum_{k=0}^{n} a_k$?

- *En posant $c_n = 1$ pour tout n, on peut écrire $b_n = \sum_{k=0}^{n} a_k c_{n-k}$. La série entière $\sum b_n z^n$ est donc la série produit de Cauchy de la série entière $\sum a_n z^n$, de rayon de convergence R, et de la série entière $\sum c_n z^n$, de rayon de convergence 1.*
  On a donc d'après le cours : $R' \geqslant \min(R,1)$.
- *Pour $|z| < \min(1,R)$ notons $S(z) = \sum_{n=0}^{+\infty} a_n z^n$ et $T(z) = \sum_{n=0}^{+\infty} b_n z^n$. D'après ce qui précède, on a : $T(z) = S(z) \sum_{n=0}^{+\infty} z^n = \frac{S(z)}{1-z}$, ce que l'on peut aussi écrire $S(z) = (1-z)T(z)$. La série entière $S(z)$ apparaît donc comme le produit de la série entière $1-z$, de rayon de convergence $+\infty$, et de la série entière $T(z)$, de rayon de convergence R'.*
  Toujours d'après le cours sur les séries entières produit, on aura $R \geqslant \min(R', +\infty) = R'$.
- *En conclusion on a : $\min(R,1) \leqslant R' \leqslant R$.*
  - *Si $R \leqslant 1$, on aura donc $R' = R$.*
  - *Si $R > 1$ on ne peut rien dire de plus que $R' \in [1\,;R]$, comme le montrent les exemples suivants.*
    - *On prend $a_n = a^n$ avec $a < 1$. Ainsi $R = \frac{1}{a} > 1$. On a alors $b_n = \sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$. Puisque $\lim_{n\to+\infty} b_n = \frac{1}{1-a}$, on aura $R' = 1$.*
    - *On prend $a_n = \frac{1}{(n-1)!} - \frac{1}{n!}$ pour $n \geqslant 1$ et $a_0 = 0$. Nous vous laissons le soin de vérifier que, dans ce cas, $R = R' = +\infty$.*

■ Comment calculer la somme d'une série entière ?

Dans tous les cas bien sûr, avant de se lancer dans le calcul de la somme d'une série entière, on commencera par déterminer son rayon de convergence !

SF13.6 Faire apparaître des séries entières connues

Il s'agit de la situation la plus simple : dans l'expression donnée, on essaye de faire apparaître des séries entières connues, par exemple à l'aide de combinaisons linéaires ou d'un changement de variable.

Exemple

Déterminer le rayon de convergence et la somme de la série entière $\sum_{n=0}^{+\infty} \operatorname{ch} n z^{2n}$.

Nous avons déjà déterminé le rayon de convergence R de cette série entière dans un exemple de **[SF13.2]** *page 512 :* $R = \dfrac{1}{\sqrt{\mathrm{e}}}$.

Pour $|z| < R$, *on peut écrire :* $\displaystyle\sum_{n=0}^{+\infty} \operatorname{ch} n z^{2n} = \frac{1}{2}\left(\sum_{n=0}^{+\infty} \mathrm{e}^{n} z^{2n} + \sum_{n=0}^{+\infty} \mathrm{e}^{-n} z^{2n}\right)$,
donc en reconnaissant des séries géométriques, la somme cherchée vaut :

$$\frac{1}{2}\left(\frac{1}{1-\mathrm{e}z^2} + \frac{1}{1-\frac{z^2}{\mathrm{e}}}\right).$$

Exemple

Déterminer le rayon de convergence et la somme de la série entière $\displaystyle\sum_{n=0}^{+\infty} (-1)^n \frac{x^n}{(2n)!}$.

La règle de d'Alembert permet aisément de montrer que le rayon de convergence de cette série entière est $+\infty$. *Notons* $f(x)$ *sa somme.*

On reconnaît « presque » le développement en série entière des fonctions cos *ou* ch. *Pour cela, on devine qu'il faut poser* $x = \pm t^2$ *; plus précisément :*

– *pour* $x \geqslant 0$, *en posant* $t = \sqrt{x}$ *on obtient :*

$$f(x) = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n}}{(2n)!} = \cos t = \cos\sqrt{x}\,;$$

– *pour* $x \geqslant 0$, *en posant* $t = \sqrt{-x}$ *on obtient :*

$$f(x) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{(2n)!} = \operatorname{ch} t = \operatorname{ch}\sqrt{-x}\,.$$

Exemple

Déterminer le rayon de convergence et la somme de la série entière $\displaystyle\sum_{n=0}^{+\infty} \frac{x^{3n}}{(3n)!}$.

La règle de d'Alembert permet aisément de montrer que le rayon de convergence de cette série entière est $+\infty$. *Notons* $f(x)$ *sa somme.*

On reconnaît ici une « partie » du développement en série entière de la fonction exp. *Pour obtenir un terme sur trois, il est légitime de penser à introduire les racines cubique de* 1 *dans* $\mathbb{C}$.

Plus précisément, en posant $\mathrm{j} = \mathrm{e}^{\frac{2\mathrm{i}\pi}{3}}$ *on sait que* $\mathrm{j}^3 = 1$ *et* $1 + \mathrm{j} + \mathrm{j}^2 = 0$ *donc pour tout entier* n *on a :*

$$1 + \mathrm{j}^n + \mathrm{j}^{2n} = \begin{cases} 3 & \text{si } n \equiv 0 \pmod 3 \\ 0 & \text{sinon} \end{cases}$$

donc $f(x) = \dfrac{1}{3}(\mathrm{e}^{x} + \mathrm{e}^{\mathrm{j}x} + \mathrm{e}^{\mathrm{j}^2 x})$.

Puisque $\mathrm{e}^{\mathrm{j}x} = \mathrm{e}^{-\frac{x}{2}+\mathrm{i}x\frac{\sqrt{3}}{2}} = \mathrm{e}^{-\frac{x}{2}}\left(\cos\left(\frac{\sqrt{3}}{2}x\right)+\mathrm{i}\sin\left(\frac{\sqrt{3}}{2}x\right)\right)$ *et que* $\mathrm{e}^{\mathrm{j}^2x} = \overline{\mathrm{e}^{\mathrm{j}x}}$ *on obtient :*

$$\forall\, x\in\mathbb{R},\ f(x) = \frac{1}{3}\left(\mathrm{e}^{x} + 2\mathrm{e}^{-\frac{x}{2}}\cos\left(\frac{\sqrt{3}}{2}x\right)\right).$$

SF13.7 Utiliser la dérivation ou l'intégration

Pour une série entière, la dérivation ou l'intégration terme à terme sur l'intervalle ouvert de convergence ont été démontrées dans le cours ; contrairement aux séries de fonctions dans le cas général, il n'y a donc rien de plus à justifier.

Exemple

Déterminer le rayon de convergence puis calculer $\displaystyle\sum_{n=1}^{+\infty} nx^n$.

La série $\displaystyle\sum_{n\in\mathbb{N}} x^n$ *a pour rayon de convergence* $R=1$*, et, pour tout* $x\in\,]-1\,;1[$*,* $f(x) = \displaystyle\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$.
Le théorème rappelé en **[S13.5]** *permet d'écrire directement, pour tout* $x\in\,]-1\,;1[$ *:* $f'(x) = \displaystyle\sum_{n=1}^{+\infty} nx^{n-1}$*, donc* $\displaystyle\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$ *et enfin :* $\displaystyle\sum_{n=1}^{+\infty} nx^n = \frac{x}{(1-x)^2}$.

✓ *Cette série entière intervient fréquemment en probabilités, il faut donc savoir retrouver rapidement sa valeur.*

Exemple

Déterminer le rayon de convergence et calculer $\displaystyle\sum_{n=1}^{+\infty} \frac{x^{4n-1}}{4n-1}$.

Il est facile de montrer que le rayon de convergence de cette série entière est égal à 1*, par l'une des nombreuses méthodes déjà exposées.*
Notons $f(x)$ *sa somme, pour* $x\in\,]-1\,;1[$*. Alors :*

$$f(x) = \sum_{n=1}^{+\infty}\int_0^x t^{4n-2}\,\mathrm{d}t = \int_0^x\left(\sum_{n=1}^{+\infty} t^{4n-2}\right)\mathrm{d}t = \int_0^x \frac{t^2}{1-t^4}\,\mathrm{d}t\,.$$

Nous n'avons pas besoin de justifier en détail ici l'interversion de $\sum$ *et* $\int$*, puisque cela est un résultat du cours sur les séries entières (***[S13.6]***).*
Or pour $t\in\,]-1\,;1[$ *:*

$$\frac{t^2}{1-t^4} = \frac{1}{4}\left(\frac{1}{1-t}+\frac{1}{1+t}-\frac{2}{1+t^2}\right),$$

ce qui donne finalement : $\forall\, x\in\,]-1\,;1[,\ f(x) = \frac{1}{4}\ln\left(\frac{1+x}{1-x}\right) - \frac{1}{2}\operatorname{Arctan} x\,.$

SF13.8 Utiliser une décomposition en éléments simples

Lorsque a_n est une fonction rationnelle de n, on pourra essayer de l'écrire comme somme d'une fonction polynôme en n et de termes de la forme $\dfrac{1}{an+b}$.

Les sommes de séries entières de la forme $\sum n^k x^n$ peuvent s'obtenir par dérivations successives de la série entière $\sum x^n$, et celles de la forme $\sum \dfrac{x^{an+b}}{an+b}$ par intégration (comme dans les exemples ci-dessus).

Exemple

Déterminer le rayon de convergence et la somme de la série entière $\displaystyle\sum_{n=1}^{+\infty} \frac{(n+1)^2}{n} x^n$.

- *Le rayon de convergence de cette série entière est égal à 1 : on peut le démontrer en utilisant le critère de d'Alembert ou en remarquant que* $\displaystyle\lim_{n\to+\infty} \frac{(n+1)^2}{n} x^n = 0 \iff |x| < 1$.
- *Pour tout* $n \in \mathbb{N}^*$ *on a* $\dfrac{(n+1)^2}{n} = n + 2 + \dfrac{1}{n}$ *donc, pour* $|x| < 1$ *:*

$$\sum_{n=1}^{+\infty} \frac{(n+1)^2}{n} x^n = \sum_{n=1}^{+\infty} n x^n + 2\sum_{n=1}^{+\infty} x^n + \sum_{n=1}^{+\infty} \frac{x^n}{n}$$

(cette écriture est licite car les trois séries écrites convergent pour $|x| < 1$*). En utilisant alors le résultat de l'exemple précédent et les développements en série entière connus, la somme cherchée vaut :*

$$\frac{x}{(1-x)^2} + \frac{2x}{1-x} - \ln(1-x) = \frac{3x - 2x^2}{(1-x)^2} - \ln(1-x)\,.$$

SF13.9 Utiliser une équation différentielle

Cette méthode peut être utilisée lorsque l'on connaît une relation de récurrence simple entre les coefficients de la série entière (voir aussi **[SF13.14]** et **[SF13.19]**). Elle consiste à trouver une équation différentielle (simple) dont la somme de la série entière cherchée est solution, puis à résoudre cette équation. La somme cherchée fait alors partie des solutions trouvées.

✓ Cette méthode est un peu délicate et demande un peu d'intuition dans les calculs, c'est pourquoi nous vous conseillons de bien étudier les exemples qui suivent et de les refaire.

Exemple

Déterminer le rayon de convergence et la somme de la série entière $\displaystyle\sum_{n=0}^{+\infty} \frac{x^{3n}}{(3n)!}$.

Cet exemple a déjà été traité de façon différente dans un exemple de **[SF13.6]** *page 516.*

Nous avons vu que le rayon de convergence de cette série entière est $+\infty$. Sa somme f est donc de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$ et pour tout x réel on a :

$$f(x)=\sum_{n=0}^{+\infty}\frac{x^{3n}}{(3n)!}\ ,\ f'(x)=\sum_{n=1}^{+\infty}\frac{x^{3n-1}}{(3n-1)!}\quad et\quad f''(x)=\sum_{n=1}^{+\infty}\frac{x^{3n-2}}{(3n-2)!}\ ,$$

de sorte que : $\forall\, x\in\mathbb{R},\ f(x)+f'(x)+f''(x)=\displaystyle\sum_{n=0}^{+\infty}\frac{x^n}{n!}=\mathrm{e}^x$.

f est donc solution de l'équation différentielle linéaire du second ordre à coefficients constants : $y''+y'+y=\mathrm{e}^x$.
La solution générale de cette équation est : $y=\dfrac{1}{3}\mathrm{e}^x+\alpha\mathrm{e}^{\mathrm{j}x}+\beta\mathrm{e}^{\mathrm{j}^2x}$ *et les conditions* $f(0)=1$, $f'(0)=0$ *permettent de déterminer la valeur de* $f(x)$.

Exemple

Pour $x\in\mathbb{R}$ on pose $f(x)=\displaystyle\sum_{n=0}^{+\infty}\frac{(-1)^n}{2n-1}\binom{2n}{n}x^n$.
Déterminer le rayon de convergence R de cette série entière et la valeur de $f(x)$.

Pour tout $n\in\mathbb{N}$ *posons* $a_n=\dfrac{(-1)^n}{2n-1}\binom{2n}{n}$.

- *Les a_n ne sont pas nuls et*

$$\frac{a_{n+1}}{a_n}=-\frac{2n-1}{2n+1}\frac{(2n+2)!}{(n+1)!^2}\frac{n!^2}{(2n)!}=-\frac{2(2n-1)}{n+1}\qquad(*)$$

Puisque $\displaystyle\lim_{n\to+\infty}\left|\frac{a_{n+1}}{a_n}\right|=4$, *le critère de d'Alembert donne* $R=\dfrac{1}{4}$.

- *Pour* $x\in\left]-\dfrac{1}{4}\,;\dfrac{1}{4}\right[$ *on a donc :*

$$f(x)=\sum_{n=0}^{+\infty}a_nx^n\quad et\quad f'(x)=\sum_{n=1}^{+\infty}na_nx^{n-1}=\sum_{n=0}^{+\infty}(n+1)a_{n+1}x^n\ .$$

La relation de récurrence $(*)$ *s'écrit aussi* $(n+1)a_{n+1}=-4na_n+2a_n$,

d'où : $\forall\, x\in\left]-\dfrac{1}{4}\,;\dfrac{1}{4}\right[,\ f'(x)=\displaystyle\sum_{n=0}^{+\infty}(-4na_n+2a_n)x^n=-4xf'(x)+2f(x)$.

f est donc solution sur $\left]-\dfrac{1}{4}\,;\dfrac{1}{4}\right[$ *de l'équation différentielle :* $(4x+1)y'=2y$. *La solution générale de cette équation est :*

$$y=C\exp\left(\int\frac{2}{4x+1}\,\mathrm{d}x\right)=C\exp\left(\frac{1}{2}\ln|4x+1|\right)=C\sqrt{4x+1}$$

et puisque $f(0)=a_0=-1$ *on obtient :* $f(x)=-\sqrt{4x+1}$.

SF13.10 Utiliser une relation entre les coefficients

Certaines relations entre les coefficients d'une série entière peuvent conduire à une équation différentielle, comme dans l'exemple précédent, ou bien permettre de trouver directement la somme.

Exemple

Calculer le rayon de convergence et la somme de la série entière $\sum_{n\geqslant 0} a_n z^n$ où :

$$a_0 = 1\,,\ a_1 = 2\,,\ a_2 = 4 \ \text{ et } \ \forall n \in \mathbb{N},\ a_{n+3} = a_{n+2} + a_{n+1} - a_n\,.$$

Soit R le rayon de convergence de la série entière $\sum_{n\geqslant 0} a_n z^n$, que l'on supposera > 0. Notons $f(z)$ la somme de cette série entière.
Pour tout z tel que $|z| < R$ on a :

$$\begin{aligned}\sum_{n=0}^{+\infty} a_{n+3} z^{n+3} &= \sum_{n=0}^{+\infty} a_{n+2} z^{n+3} + \sum_{n=0}^{+\infty} a_{n+1} z^{n+3} - \sum_{n=0}^{+\infty} a_n z^{n+3} \\ &= \sum_{n=2}^{+\infty} a_n z^{n+1} + \sum_{n=1}^{+\infty} a_n z^{n+2} - z^3 \sum_{n=0}^{+\infty} a_n z^n\,,\end{aligned}$$

d'où $f(z) - a_2 z^2 - a_1 z - a_0 = z\big(f(z) - a_1 z - a_0\big) + z^2\big(f(z) - a_0\big) - z^3 f(z)$,
soit, compte tenu des valeurs initiales :

$$\underbrace{(1 - z - z^2 + z^3)}_{=(1-z)^2(1+z)} f(z) = 1 + z + z^2\,.$$

On en déduit $R = 1$ et pour tout z tel que $|z| < 1$: $f(z) = \dfrac{1 + z + z^2}{(1-z)^2(1+z)}$.

■ Comment trouver le développement en série entière d'une fonction ?

SF13.11 Utiliser la série de Taylor

La série de Taylor d'une application f d'une variable réelle de classe $\mathscr{C}^\infty$ au voisinage de 0 est $\sum_{n\geqslant 0} \dfrac{f^{(n)}(0)}{n!} x^n$.

Si la série de Taylor de f converge *vers* f, alors ce sera le développement en série entière de f.

Le problème de cette méthode est donc de vérifier ce point, par exemple en majorant le reste dans la formule de Taylor (reste intégrale ou inégalité de Taylor-Lagrange). C'est une méthode assez restrictive, à n'utiliser que si l'on sait calculer les dérivées successives, et si elles sont simples.
Cette méthode a surtout servi en cours pour obtenir les premiers développements en série entière des fonctions usuelles (exp, sin, cos…)

Exemple

Soit $a \in]-1\,;1[$ et $f\colon x \mapsto \sum_{n=0}^{+\infty} \sin(a^n x)$.

Déterminer le domaine de définition de f et en déterminer le développement en série entière.

On notera $u_n(x) = \sin(a^n x)$ pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$.

- *On connaît la majoration :* $\forall x \in \mathbb{R},\ |\sin x| \leqslant |x|$. *On a donc* $|u_n(x)| \leqslant |a|^n |x|$ *et puisque la série géométrique* $\sum |a|^n$ *converge, on en déduit que la série* $\sum u_n(x)$ *est absolument convergente. Donc* f *est définie sur* $\mathbb{R}$.
- *Montrons que* f *est de classe* $\mathscr{C}^\infty$ *sur* $\mathbb{R}$, *en utilisant le théorème sur la dérivation itérée d'une série de fonctions* **[S12.11]**.
  - *Les* u_n *sont de classe* $\mathscr{C}^\infty$ *sur* $\mathbb{R}$ *et, pour tout* $k \in \mathbb{N}$ *et tout* $x \in \mathbb{R}$ *on a* $u_n^{(k)}(x) = a^{nk} \sin\left(a^n x + k\frac{\pi}{2}\right)$.
  - *Pour tout* $k \in \mathbb{N}^*$ *on aura donc* $\left\|u_n^{(k)}\right\|_\infty^{\mathbb{R}} \leqslant |a|^{nk}$ *qui est le terme d'une série géométrique de raison* $|a|^k < 1$ *donc convergente. Les séries* $\sum_{n\in\mathbb{N}} u_n^{(k)}$ *sont donc toutes normalement donc uniformément convergentes sur* $\mathbb{R}$.

  Cela suffit pour affirmer que f *est de classe* $\mathscr{C}^\infty$ *sur* $\mathbb{R}$ *et, pour tout* $k \in \mathbb{N}$ *on a* $f^{(k)}(x) = \sum_{n=0}^{+\infty} a^{nk} \sin\left(a^n x + k\frac{\pi}{2}\right)$.
- *De la formule précédente on tire :*

$$\forall k \in \mathbb{N},\ f^{(2k)}(0) = 0 \quad et \quad f^{(2k+1)}(0) = \sum_{n=0}^{+\infty} (-1)^k a^{(2k+1)n} = \frac{(-1)^k}{1-a^{2k+1}}\cdot$$

De plus, pour tout $k \in \mathbb{N}^*$ *et tout* $x \in \mathbb{R}$ *on a :*

$$\left|f^{(k)}(x)\right| \leqslant \sum_{n=0}^{+\infty} \left|a^{nk} \sin\left(a^n x + k\frac{\pi}{2}\right)\right| \leqslant \sum_{n=0}^{+\infty} |a|^{nk} = \frac{1}{1-|a|^k},$$

donc l'inégalité de Taylor-Lagrange donne :

$$\left|f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k\right| \leqslant \left\|f^{(n+1)}\right\|_\infty \frac{|x|^{n+1}}{(n+1)!}$$
$$\leqslant \frac{|x|^{n+1}}{(n+1)!(1-|a|^{n+1})} \underset{n\to+\infty}{\longrightarrow} 0$$

ce qui prouve que la série de Taylor de f *converge simplement vers* f.

En conclusion, pour tout $x \in \mathbb{R}$ *on a :*

$$f(x) = \sum_{k=0}^{+\infty} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{+\infty} \frac{(-1)^k}{1-a^{2k+1}} \frac{x^{2k+1}}{(2k+1)!}\cdot$$

✓ *Une autre solution était possible : elle consiste à écrire le développement en série entière de chaque fonction* $x \mapsto \sin(a^n x)$, *puis à intervertir les signes* $\sum$, *en justifiant cette interversion à l'aide du théorème de Fubini ; un exemple d'utilisation de cette méthode a été donné en* **[SF11.24]** *page 475 ; voir aussi* **[SF13.15]**).

SF13.12 Utiliser les développements en série entière usuels

Cette méthode consiste à écrire le développement en série entière des fonctions qui figurent dans l'énoncé, puis à utiliser les opérations telles que combinaisons linéaires, produit, dérivation, intégration.

> ✓ La dérivation des fonctions d'une variable complexe n'étant pas au programme, les méthodes de dérivation et d'intégration sont à réserver aux fonctions d'une variable réelle.

Exemple

Développer en série entière $f : x \mapsto \ln(1+x+x^2)$.

On remarque déjà que f est bien définie sur $\mathbb{R}$ puisque $1+x+x^2>0$ pour tout x.
Pour $x \in]-1\,;1[$ on a $f(x) = \ln\left(\dfrac{1-x^3}{1-x}\right)$ donc en utilisant le développement en série entière de la fonction ln *:*

$$\forall\, x \in]-1\,;1[,\ f(x) = \ln(1-x^3) - \ln(1-x) = -\sum_{n=1}^{+\infty} \frac{x^{3n}}{n} + \sum_{n=1}^{+\infty} \frac{x^n}{n}$$

$$\text{soit } f(x) = \sum_{n=1}^{+\infty} a_n x^n \text{ avec } a_n = \begin{cases} \dfrac{1}{n} & \text{si } n \not\equiv 0 \pmod 3 \\ -\dfrac{2}{n} & \text{si } n \equiv 0 \pmod 3 . \end{cases}$$

Le rayon de convergence de cette série entière est égal à 1.

Exemple

Montrer que, pour tout $z \in \mathbb{C}$ tel que $|z|<1$ et pour tout $p \in \mathbb{N}$, on a :

$$\frac{1}{(1-z)^{p+1}} = \sum_{n=0}^{+\infty} \binom{n+p}{p} z^n.$$

L'égalité demandée peut se démontrer de deux façons, les deux étant instructives.

- Première solution : par récurrence sur p
  La formule est exacte pour $p=0$ (série géométrique).
  Supposons que, à un certain rang p, l'on ait $\dfrac{1}{(1-z)^{p+1}} = \displaystyle\sum_{n=0}^{+\infty} \binom{n+p}{p} z^n$.
  En effectuant le produit de Cauchy de cette série par celle donnant $\dfrac{1}{1-z}$, *toutes deux ayant pour rayon de convergence* 1, *on obtient :*

$$\forall\, z \in \mathbb{C} \text{ tel que } |z|<1,\ \frac{1}{(1-z)^{p+2}} = \sum_{n=0}^{+\infty} a_n z^n$$

  avec $a_n = \displaystyle\sum_{k=0}^{n} \binom{k+p}{p} = \sum_{k=0}^{n} \binom{k+p}{k}$.
  Or la formule du triangle de Pascal donne

$$\binom{k+1+p}{k} = \binom{k+p}{k} + \binom{k+p}{k-1},$$

avec la convention habituelle $\binom{N}{k} = 0$ *si* $k < 0$, *d'où :*

$$a_n = \sum_{k=0}^{n} \left[\binom{k+1+p}{k} - \binom{k+p}{k-1} \right] = \binom{n+p+1}{n} = \binom{n+p+1}{p+1}$$

(télescopage), ce qui donne bien la première formule à l'ordre $p+1$.

- Deuxième solution : par dérivation

  Soit z *non nul tel que* $|z| < 1$. *Pour tout* réel t *tel que* $t \in \left]-\frac{1}{|z|} ; \frac{1}{|z|}\right[$, *on pose* $f(t) = \frac{1}{1-tz}$.

  f *est évidemment de classe* $\mathscr{C}^\infty$ *sur* $\left]-\frac{1}{|z|} ; \frac{1}{|z|}\right[$, *et pour tout entier* p, *on a* $f^{(p)}(t) = \frac{p!z^p}{(1-tz)^{p+1}}$.

  D'autre part, puisque $|tz| < 1$, *on a le développement en série entière* $f(t) = \sum_{n=0}^{+\infty} (tz)^n$, *de rayon de convergence* $\frac{1}{|z|}$.

  On peut donc la dériver terme à terme p *fois sur son intervalle ouvert de convergence et l'on obtient :*

  $$\forall p \in \mathbb{N}\ ,\ \forall t \in \left]-\frac{1}{|z|} ; \frac{1}{|z|}\right[\ ,\ f^{(p)}(t) = \sum_{n=p}^{+\infty} n(n-1)\cdots(n-p+1)t^{n-p}z^n\ ,$$

  soit finalement :

  $$\frac{p!z^p}{(1-tz)^{p+1}} = \sum_{n=p}^{+\infty} n(n-1)\cdots(n-p+1)t^{n-p}z^n.$$

  Puisque $|z| < 1$, $t = 1$ *appartient bien à l'intervalle* $\left]-\frac{1}{|z|} ; \frac{1}{|z|}\right[$, *et pour* $t = 1$, *on obtient bien la formule voulue.*

  ✓ Dans le cas où z est réel, une autre façon de calculer le développement en série entière de $\frac{1}{(1-x)^{p+1}} = (1-x)^{-p-1}$ consiste à utiliser directement le développement en série entière de $(1+x)^\alpha$ vu en cours. C'est d'ailleurs la méthode la plus rapide dans ce cas, et nous vous invitons à refaire les calculs de cette façon.

  ✓ Ce développement en série entière est utilisé dans de nombreux exercices de probabilités. Il peut d'ailleurs se démontrer par des méthodes probabilistes.

SF13.13 Développer en série entière une fraction rationnelle

- Le développement en série entière d'une fraction rationnelle, sous réserve qu'elle soit définie au voisinage de 0 c'est-à-dire n'admette pas 0 pour pôle, se fait en se ramenant aux développements en série entière vus ci-dessus. Le principe de la méthode est simple, même si le calcul est parfois un peu long.

Soit $F \in \mathbb{C}(X)$ une fraction rationnelle n'admettant pas 0 pour pôle.
On écrit d'abord la décomposition en éléments simples de F dans $\mathbb{C}(X)$:

$$F = E \ + \sum_{i=1}^{n} \sum_{r=1}^{r_i} \frac{\lambda_{i,r}}{(X-a_i)^r}$$

où E est un polynôme, les $\lambda_{i,r}$ des nombres complexes, et où les a_i sont les pôles de F, distincts deux à deux et non nuls.
Soit $R = \min_{i \in [\![1;n]\!]} |a_i|$ $(R>0)$; pour $z \in \mathbb{C}$ tel que $|z| < R$, on a :

$$\begin{aligned} F(z) &= E(z) \ + \sum_{i=1}^{n} \sum_{r=1}^{r_i} \frac{\lambda_{i,r}}{(z-a_i)^r} \\ &= E(z) \ + \sum_{i=1}^{n} \sum_{r=1}^{r_i} \frac{\lambda_{i,r}}{(-a_i)^r} \cdot \frac{1}{\left(1-\frac{z}{a_i}\right)^r} \\ &= E(z) \ + \sum_{i=1}^{n} \sum_{r=1}^{r_i} \frac{\lambda_{i,r}}{(-a_i)^r} \sum_{k=0}^{+\infty} \binom{k+r-1}{r-1} \left(\frac{z}{a_i}\right)^k \end{aligned}$$

ce qui, après réarrangement des $\sum$ donne bien le développement en série entière de $F(z)$.

> ✓ *On voit ici l'utilité du développement en série entière vu juste au-dessus, page 522.*

Ce développement a, a priori, un rayon de convergence $\geqslant R$; mais ce rayon est en fait exactement égal à R car il existe un des pôles a_{i_0} tel que $|a_{i_0}| = R$, et F n'est pas définie en ce point.

- Dans le cas d'une fraction rationnelle de $\mathbb{R}(X)$, le plus simple en général est de la décomposer en éléments simples dans $\mathbb{C}(X)$ puis de regrouper convenablement les termes à la fin.

Exemple

Donner le développement en série entière de $\dfrac{1}{(z-1)(z^2+1)}$.

Notons $f(z) = \dfrac{1}{(z-1)(z^2+1)}$.
f est définie sur $\mathbb{C} \setminus \{\pm \mathrm{i}, 1\}$ et pour tout z de cet ensemble on a :

$$\begin{aligned} f(z) &= \frac{1}{2(z-1)} - \frac{1-\mathrm{i}}{4(z-\mathrm{i})} - \frac{1+\mathrm{i}}{4(z+\mathrm{i})} \\ &= -\frac{1}{2}\frac{1}{1-z} - \frac{1+\mathrm{i}}{4}\frac{1}{1+\mathrm{i}z} - \frac{1-\mathrm{i}}{4}\frac{1}{1-\mathrm{i}z}. \end{aligned}$$

Pour tout z tel que $|z| < 1$ on aura donc le développement en série entière :

$$f(z) = -\frac{1}{2}\sum_{n=0}^{+\infty} z^n - \frac{1+\mathrm{i}}{4}\sum_{n=0}^{+\infty} (-\mathrm{i})^n z^n - \frac{1-\mathrm{i}}{4}\sum_{n=0}^{+\infty} \mathrm{i}^n z^n = \sum_{n=0}^{+\infty} a_n z^n$$

$$\text{avec } a_n = -\frac{1}{2} - \mathrm{i}^n\left(\frac{1-\mathrm{i}}{4} + (-1)^n \frac{1+\mathrm{i}}{4}\right) = \begin{cases} -1 & \text{si } n \equiv 0 \pmod 4 \\ -1 & \text{si } n \equiv 1 \pmod 4 \\ 0 & \text{si } n \equiv 2 \pmod 4 \\ 0 & \text{si } n \equiv 3 \pmod 4 \end{cases}.$$

SF13.14 Utiliser une équation différentielle

Cette méthode consiste à « injecter » l'expression de f sous forme d'une série entière $f(x) = \sum_{n=0}^{+\infty} a_n x^n$, ainsi que celle de ses dérivées, dans une équation différentielle vérifiée par f pour obtenir une relation de récurrence entre les coefficients, puis les calculer.

Le problème de cette méthode est qu'elle suppose f développable en série entière *a priori*, donc il faut vérifier a posteriori que le résultat obtenu convient, c'est-à-dire que le rayon de convergence de la série obtenue est bien strictement positif, à moins que l'on ait pu prouver à l'avance par les théorèmes **[S13.2]** que f est effectivement développable en série entière.

Exemple

Soit $\alpha \in \mathbb{R}^*$. Développer en série entière $f : x \mapsto \left(x + \sqrt{1+x^2}\right)^\alpha$.

- *On remarque déjà que* $1 + x^2 > 0$ *pour tout* $x \in \mathbb{R}$ *et que, puisque* $\sqrt{1+x^2} > \sqrt{x^2} = |x|$, *on aura aussi* $x + \sqrt{1+x^2} > 0$ *pour tout* x, *donc* f *est bien définie sur* $\mathbb{R}$. *Elle y est aussi de classe* $\mathscr{C}^\infty$ *d'après les théorèmes usuels.*
- *Un calcul rapide montre que pour tout* x *réel :*

$$f'(x) = \alpha\left(1 + \frac{x}{\sqrt{1+x^2}}\right)\left(x + \sqrt{1+x^2}\right)^{\alpha-1} = \frac{\alpha}{\sqrt{1+x^2}} f(x)$$

puis

$$f''(x) = -\frac{\alpha x}{(1+x^2)^{3/2}} f(x) + \frac{\alpha}{\sqrt{1+x^2}} f'(x) = -\frac{x}{1+x^2} f'(x) + \frac{\alpha^2}{1+x^2} f(x)\,.$$

Ainsi f *est solution de l'équation différentielle* $(1+x^2)y'' + xy - \alpha^2 y = 0$.

- *En supposant* a priori f *développable en série entière sur un intervalle* $]-R\,;R[$ *avec* $R > 0$, *on pourra écrire, pour tout* $|x| < R$ *:*

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n \ , \ f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \ , \ f''(x) = \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}$$

d'où en remplaçant dans l'équation différentielle :

$$\sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2} + \sum_{\substack{n=2 \\ n=0}}^{+\infty} n(n-1)a_n x^n + \sum_{\substack{n=1 \\ n=0}}^{+\infty} n a_n x^n - \alpha^2 \sum_{n=0}^{+\infty} a_n x^n = 0\,.$$

> ✓ *Vous noterez l'astuce consistant à faire partir la seconde somme à* $n = 0$, *ce qui est possible car les termes d'indices* $n = 0$ *et* $n = 1$ *sont nuls (même principe pour la troisième somme) ; cela permet de faire débuter les sommes à partir du même indice.*

Après un décalage d'indices dans la première somme on obtient :

$$\sum_{n=0}^{+\infty} \left[(n+2)(n+1)a_{n+2} + \left(n(n-1) + n - \alpha^2\right) a_n\right] x^n = 0\,.$$

Par unicité du développement en série entière, on en déduit la relation de récurrence :

$$\forall\, n \in \mathbb{N},\ a_{n+2} = \frac{\alpha^2 - n^2}{(n+1)(n+2)}\, a_n\,.$$

Compte tenu de $a_0 = f(0) = 1$ *et de* $a_1 = f'(0) = \alpha$, *on en déduit, pour* $n \in \mathbb{N}^*$ *:*

$$a_{2n+2} = \frac{\alpha^2(\alpha^2-4)\cdots(\alpha^2-4n^2)}{(2n)!} \quad \text{et}$$

$$a_{2n+1} = \frac{\alpha(\alpha^2-1)\cdots(\alpha^2-(2n-1)^2)}{(2n-1)!}\,.$$

- *Il reste à faire la synthèse des calculs précédents.*
  Soit donc $(a_n)_{n\in\mathbb{N}}$ *donnée par les relations ci-dessus. On vérifie alors que la série entière* $\sum a_n x^n$ *a pour rayon de convergence* 1 *(cela se démontre par exemple en utilisant la règle de d'Alembert).*
  Notons $S(x) = \sum_{n=0}^{+\infty} a_n x^n$. *Alors* S *est de classe* $\mathscr{C}^\infty$ *sur* $]-1\,;1[$, *est solution de l'équation différentielle* $(1+x^2)y'' + xy - \alpha^2 y = 0$ *(d'après les calculs faits), et vérifie* $S(0) = 1$, $S'(0) = \alpha$ *; d'après le théorème de Cauchy-Lipschitz linéaire (voir* **[S17.7]***) elle coïncide avec* f *sur* $]-1\,;1[$.
  La série entière $\sum a_n x^n$ *est donc bien le développement en série entière de* f *sur* $]-1\,;1[$.

SF13.15 Utiliser le théorème de Fubini

L'utilisation du théorème de Fubini pour trouver un développement en série entière est presque incontournable lorsque la fonction considérée est définie comme somme d'une série de fonctions.

Exemple

Soit $(a_n)_{n\in\mathbb{N}}$ une suite complexe telles que $\sum |a_n|^2$ converge.

Montrer que la fonction $f\colon x \mapsto \sum_{n=1}^{+\infty} \frac{a_n}{n-x}$ est développable en série entière sur $]-1\,;1[$, et donner ce développement.

- *Démontrons d'abord la convergence de la série définissant* f.
  Pour cela, on utilise l'inégalité bien connue :

$$\forall\,(a,b) \in \mathbb{R}^2,\ ab \leqslant \frac{1}{2}(a^2+b^2)\,.$$

  On a donc pour tout $n \geqslant 1$ *et tout* $x \in\,]-1\,;1[$ *:*

$$\left|\frac{a_n}{n-x}\right| \leqslant \frac{1}{2}\left(|a_n|^2 + \frac{1}{(n-x)^2}\right),$$

  et puisque les séries de termes généraux $|a_n|^2$ *et* $\frac{1}{(n-x)^2}$ *sont convergentes, la série* $\sum_{n\geqslant 1} \frac{a_n}{n-x}$ *est absolument convergente.*

- *Pour tout $n \in \mathbb{N}^*$ et tout $x \in]-1\,;1[$ on a, puisque $\left|\frac{x}{n}\right| < 1$:*

$$\frac{1}{n-x} = \frac{1}{n\left(1-\frac{x}{n}\right)} = \frac{1}{n}\sum_{k=0}^{+\infty}\frac{x^k}{n^k}\,.$$

On peut donc écrire, pour $x \in]-1\,;1[$:

$$f(x) = \sum_{n=1}^{+\infty}\sum_{k=0}^{+\infty} u_{n,k} \quad \text{avec} \quad u_{n,k} = \frac{a_n x^k}{n^{k+1}}\,.$$

Vérifions les hypothèses du théorème de Fubini :

– *Pour tout $n \in \mathbb{N}^*$, la série $\sum_{k\in\mathbb{N}} |u_{n,k}|$ est convergente ; elle a pour somme $s_n = \dfrac{|a_n|}{n-|x|}$.*

– *La série $\sum_{n\in\mathbb{N}^*} s_n$ converge, puisque l'on a vu que la série qui définit $f(|x|)$ est absolument convergente.*

De ce théorème, on déduit que la suite double $(u_{n,k})$ est sommable et que l'on peut intervertir l'ordre des sommations c'est-à-dire :

$$\forall\, x \in]-1\,;1[,\ f(x) = \sum_{k=0}^{+\infty}\sum_{n=1}^{+\infty} u_{n,k} = \sum_{k=0}^{+\infty}\left(\sum_{n=1}^{+\infty}\frac{a_n}{n^{k+1}}\right)x^k\,.$$

On a donc obtenu le développement en série entière de f.

■ Comment utiliser les développements en série entière ?

SF13.16 Montrer qu'une fonction est de classe $\mathscr{C}^\infty$

Pour montrer qu'une fonction est de classe $\mathscr{C}^\infty$ sur un intervalle de la forme $]-R\,;R[$, on peut essayer d'en obtenir un développement en série entière.

Exemple

Soit $f : \mathbb{R}^* \longrightarrow \mathbb{R}$ définie par $f(x) = \dfrac{\mathrm{e}^x - 1}{x}$.

Montrer que f se prolonge en une fonction de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$.

Puisque $\mathrm{e}^x - 1 \underset{x\to 0}{\sim} x$, on a $\lim\limits_{x\to 0} f(x) = 1$. On pourra donc prolonger f en posant $f(0) = 1$.
Le développement en série entière de l'exponentielle donne alors

$$\forall\, x \neq 0,\ f(x) = \frac{\left(\sum\limits_{n=0}^{+\infty}\frac{x^n}{n!}\right) - 1}{x} = \frac{\sum\limits_{n=1}^{+\infty}\frac{x^n}{n!}}{x} = \sum_{n=1}^{+\infty}\frac{x^{n-1}}{n!} = \sum_{n=0}^{+\infty}\frac{x^n}{(n+1)!}\,,$$

égalité qui reste valable pour $x = 0$.
f est donc la somme sur $\mathbb{R}$ d'une série entière de rayon de convergence $+\infty$, donc est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$ d'après **[S13.5]**.

SF13.17 Calculer la somme d'une série numérique

Pour calculer la somme d'une série de la forme $\sum a_n x_0^n$ (après avoir démontré sa convergence !), on peut introduire la série entière $\sum a_n x^n$. Puisque la série numérique $\sum a_n x_0^n$ converge, le rayon de convergence de cette série entière est supérieur ou égal à $|x_0|$.
On peut alors essayer de calculer sa somme $S(x)$ pour $|x| < |x_0|$, puis examiner ce qui se passe lorsque $x \to x_0$.

Exemple

Calculer : $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{(4n+2)2^n}$ et $\displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n}{4n+2}$.

- *Nous considérons ici la série entière* $f(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{x^{4n+2}}{4n+2}$. *Il est facile de vérifier que le rayon de convergence de cette série entière est égal à* 1.
f est donc de classe $\mathscr{C}^\infty$ sur $]-1\,;1[$ et d'après le théorème sur la dérivation d'une série entière :

$$\forall\, x \in]-1\,;1[,\ f'(x) = \sum_{n=0}^{+\infty} x^{4n+1} = \frac{x}{1-x^4} \quad \textit{(série géométrique).}$$

Puisque $f(0) = 0$ on a :

$$\forall\, x \in]-1\,;1[,\ f(x) = \int_0^x \frac{t}{1-t^4}\,\mathrm{d}t = \frac{1}{2}\int_0^{x^2} \frac{\mathrm{d}u}{1-u^2} = \frac{1}{4}\ln\left(\frac{1+x^2}{1-x^2}\right).$$

On a donc, pour $x = \dfrac{1}{2^{1/4}}$: $f\left(\dfrac{1}{2^{1/4}}\right) = \displaystyle\sum_{n=0}^{+\infty} \frac{1}{(4n+2)2^{n+1/2}}$ *et finalement :*

$$\sum_{n=0}^{+\infty} \frac{1}{(4n+2)2^n} = \sqrt{2} f\left(\frac{1}{2^{1/4}}\right) = \frac{\sqrt{2}}{4}\ln\left(\frac{1+\sqrt{2}}{\sqrt{2}-1}\right) = \frac{\sqrt{2}}{2}\ln(1+\sqrt{2}).$$

- *Pour le calcul de la deuxième somme, nous ne pouvons pas utiliser la même fonction (puisqu'il faudrait considérer x réel tel que $x^4 = -1\ldots$).*
Nous allons donc considérer la série entière $g(x) = \displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n x^{4n+2}}{4n+2}$. *Le rayon de convergence de cette série entière est égal à* 1. *De plus, pour $x \in [-1\,;1]$, la suite $n \mapsto \dfrac{(-1)^n x^{4n+2}}{4n+2}$ vérifie les hypothèses du critère spécial sur les séries alternées. On peut donc majorer le reste R_n d'ordre n de cette série par :*

$$\forall\, x \in [-1\,;1],\ |R_n(x)| \leqslant \frac{x^{4n+6}}{4n+6} \leqslant \frac{1}{4n+6}$$

donc $\lim\limits_{n\to+\infty} \|R_n\|_\infty^{[-1;1]} = 0$. *Cela prouve que la série de fonctions converge normalement, donc uniformément, sur $[-1\,;1]$, donc g est continue sur $[-1\,;1]$ En particulier, la somme cherchée est égale à* $g(1) = \lim\limits_{x\to1^-} g(x)$.

Il ne reste plus qu'à calculer $g(x)$. Par la même méthode que dans l'exemple précédent (que nous ne détaillons pas) on obtient :

$$\forall\, x \in]-1\,;1[,\ g(x) = \int_0^x \frac{t}{1+t^4}\,\mathrm{d}t = \frac{1}{2}\int_0^{x^2} \frac{\mathrm{d}u}{1+u^2} = \frac{1}{2}\operatorname{Arctan}(x^2)\,.$$

Finalement : $\displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n}{4n+2} = g(1) = \frac{1}{2}\operatorname{Arctan} 1 = \frac{\pi}{8}\cdot$

SF13.18 Calculer une intégrale

L'idée est ici de développer en série entière la fonction à intégrer, puis de justifier que l'on peut intervertir série et intégrale à l'aide de l'un des théorèmes du cours **[S12.10]** ou **[S14.10]**.

Exemple

Pour tout entier $n \in \mathbb{N}$, calculer les intégrales

$$I_n = \int_0^{2\pi} \mathrm{e}^{\cos t}\cos(nt - \sin t)\,\mathrm{d}t \quad \text{et} \quad J_n = \int_0^{2\pi} \mathrm{e}^{\cos t}\sin(nt - \sin t)\,\mathrm{d}t\,.$$

- *On remarque déjà que les intégrales proposées existent bien, puisque les fonctions à intégrer sont continues sur le segment* $[0\,;2\pi]$.

  ✓ *Ce doit être un réflexe : lorsqu'un énoncé fait intervenir une intégrale, il faut penser à justifier son existence.*

- *On pense naturellement à utiliser les nombres complexes, en considérant*

$$K_n = I_n + \mathrm{i}J_n = \int_0^{2\pi} \mathrm{e}^{\cos t}\mathrm{e}^{\mathrm{i}(nt-\sin t)}\,\mathrm{d}t = \int_0^{2\pi} \mathrm{e}^{\mathrm{e}^{-\mathrm{i}t}}\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t\,.$$

  À l'aide du développement en série entière de e^z*, de rayon de convergence infini, on obtient :*

$$K_n = \int_0^{2\pi} \left(\sum_{k=0}^{+\infty} \frac{\left(\mathrm{e}^{-\mathrm{i}t}\right)^k}{k!}\right)\mathrm{e}^{\mathrm{i}nt}\,\mathrm{d}t \quad (*)$$

- *Posons alors, pour tout* $k \in \mathbb{N}$ *et tout* $t \in [0\,;2\pi]$ *:*

$$u_k(t) = \frac{\left(\mathrm{e}^{-\mathrm{i}t}\right)^k}{k!}\mathrm{e}^{\mathrm{i}nt} = \frac{\mathrm{e}^{\mathrm{i}(n-k)t}}{k!}\,.$$

  On vérifie alors les hypothèses du théorème d'interversion **[S12.10]**.

  – *Les fonctions* u_k *sont continues sur* $[0\,;2\pi]$.

  – *Pour tout* $k \in \mathbb{N}$, $\|u_k\|_\infty^{[0;2\pi]} = \dfrac{1}{k!}$, *qui est le terme général d'une série convergente, donc la série de fonctions* $\sum u_k$ *converge normalement donc uniformément sur le segment* $[0\,;2\pi]$.

  On peut donc permuter $\sum$ *et* $\int$ *dans* $(*)$ *ce qui donne :*

$$K_n = \sum_{k=0}^{+\infty} \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}(n-k)t}}{k!}\,\mathrm{d}t\,.$$

Un calcul simple donne : $\displaystyle\int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}(n-k)t}}{k!}\,\mathrm{d}t = \begin{cases} 0 & \text{si } k \neq n \\ \dfrac{2\pi}{n!} & \text{si } k = n \end{cases}$, *donc finalement :*

$$K_n = \frac{2\pi}{n!} \quad \text{puis} \quad I_n = \frac{2\pi}{n!} \text{ et } J_n = 0\,.$$

SF13.19 Calculer les termes d'une suite donnée par une relation de récurrence

Étant donnée une suite $(a_n)_{n\in\mathbb{N}}$, cette méthode consiste à considérer la fonction génératrice de cette suite, donnée par $G_a(x) = \displaystyle\sum_{n=0}^{+\infty} a_n x^n$ lorsque cette série converge. Si l'on sait calculer G_a, on pourra en déduire les coefficients a_n, soit en faisant le développement de G_a en série entière, soit en calculant ses dérivées n^{es}.

✓ Notez le lien entre cette notion et celle de fonction génératrice d'une variable aléatoire à valeurs entières.

Exemple

Soit $(a_n)_{n\in\mathbb{N}}$ une suite réelle telle que $a_1 \geqslant a_0 > 0$ et

$$\forall\, n \geqslant 2,\ a_n = a_{n-1} + 2\frac{a_{n-2}}{n}\,.$$

1) Étudier la convergence de cette suite.

2) Montrer que la suite $\left(\dfrac{a_n}{n^2}\right)_{n\in\mathbb{N}^*}$ converge.

3) a) En déduire le rayon de convergence de la série entière $\displaystyle\sum_{n\in\mathbb{N}} a_n x^n$.

b) Dans le cas $a_0 = a_1 = 1$, calculer la somme de cette série entière en utilisant une équation différentielle linéaire du premier ordre.

c) En déduire une expression de a_n en fonction de n.

1) *Par récurrence (double) immédiate on a, pour tout entier n, $a_n > 0$, et donc la suite $(a_n)_{n\in\mathbb{N}}$ est croissante. Par suite, on a*

$$\forall\, n \geqslant 2,\ a_n - a_{n-1} \geqslant \frac{2a_0}{n}\,,$$

d'où en sommant ces relations :

$$\forall\, n \geqslant 2,\ a_n - a_1 \geqslant 2a_0 \sum_{k=2}^{n} \frac{1}{k} \underset{n\to+\infty}{\longrightarrow} +\infty\,,$$

et donc la suite $(a_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$.

2) *On a pour tout $n \geqslant 2$, compte tenu de la croissance de la suite $(a_n)_{n\in\mathbb{N}}$:*

$$\begin{aligned} \frac{a_n}{n^2} - \frac{a_{n-1}}{(n-1)^2} &= a_{n-1}\left(\frac{1}{n^2} - \frac{1}{(n-1)^2}\right) + \frac{2a_{n-2}}{n^3} \\ &\leqslant a_{n-1}\left(\frac{1}{n^2} - \frac{1}{(n-1)^2} + \frac{2}{n^3}\right) = a_{n-1}\frac{2-3n}{n^3(n-1)^2} \leqslant 0\,. \end{aligned}$$

La suite $\left(\dfrac{a_n}{n^2}\right)_{n\in\mathbb{N}^*}$ *est donc décroissante, et comme elle est minorée par* 0 *on en conclut qu'elle converge. Notons* ℓ *sa limite.*

3) a) *D'après la question précédente, et puisque la série entière* $\sum n^2x^n$ *a pour rayon de convergence* 1 *:*

– *soit* $\ell \neq 0$*, et alors* $a_n \sim \ell n^2$ *donc directement* $R = 1$ *;*

– *soit* $\ell = 0$ *et alors* $a_n = o\big(n^2\big)$ *donc* $R \geqslant 1$ *mais puisque* $\sum a_n$ *diverge on a aussi* $R \leqslant 1$ *donc* $R = 1$ *encore.*

b) *Notons* $f(x) = \displaystyle\sum_{n=0}^{+\infty} a_n x^n$ *pour* $|x| < 1$*. Par dérivation d'une série entière :* $f'(x) = \displaystyle\sum_{n=1}^{+\infty} na_n x^{n-1}$*. Pour* $n \geqslant 2$ *on a :*

$$na_n = (n-1)a_{n-1} + a_{n-1} + 2a_{n-2}\,,$$

$$\begin{aligned} \text{d'où :}\ \sum_{n=2}^{+\infty} na_n x^n &= \sum_{n=2}^{+\infty}(n-1)a_{n-1}x^n + \sum_{n=2}^{+\infty} a_{n-1}x^n + 2\sum_{n=2}^{+\infty} a_{n-2}x^n \\ &= \sum_{n=1}^{+\infty} na_n x^{n+1} + x\sum_{n=1}^{+\infty} a_n x^n + 2x^2\sum_{n=0}^{+\infty} a_n x^n \end{aligned}$$

soit $xf'(x) - a_1 x = x^2 f'(x) + x(f(x) - a_0) + 2x^2 f(x)$*.*

Compte tenu de $a_0 = a_1 = 1$ *on obtient :*

$$x(1-x)f'(x) = x(2x+1)f(x)\,.$$

On a donc $(1-x)f'(x) = (2x+1)f(x)$ *pour* $x \neq 0$ *et cette égalité reste vraie pour* $x = 0$ *par continuité.*

Il n'y a plus qu'à résoudre l'équation différentielle $(1-x)y' = (2x+1)y$ *sur* $]-1\,;1[$*; c'est une équation linéaire homogène du premier ordre dont les solutions sont données par :*

$$\begin{aligned} y &= C\exp\left(\int \frac{2x+1}{1-x}\,\mathrm{d}x\right) = C\exp\left(\int \frac{2x-2+3}{1-x}\,\mathrm{d}x\right) \\ &= C\exp\left(\int\left(-2+\frac{3}{1-x}\right)\mathrm{d}x\right) = C\frac{\mathrm{e}^{-2x}}{(1-x)^3}\cdot \end{aligned}$$

La condition $f(0) = 1$ *donne alors* $f(x) = \dfrac{\mathrm{e}^{-2x}}{(1-x)^3}\cdot$

c) *On connaît le développement en série entière :*

$$\forall\, x \in \mathbb{R},\ \mathrm{e}^{-2x} = \sum_{n=0}^{+\infty} \frac{(-1)^n 2^n}{n!}x^n$$

et, en dérivant deux fois le développement en série entière de $\dfrac{1}{1-x}$ *:*

$$\forall\, x \in\,]-1\,;1[,\ \frac{1}{(1-x)^3} = \frac{1}{2}\sum_{n=2}^{+\infty} n(n-1)x^{n-2} = \frac{1}{2}\sum_{n=0}^{+\infty}(n+2)(n+1)x^n.$$

Le développement en série entière de f est le produit de Cauchy de ces deux séries entières, donc avec les formules du cours on obtient :

$$a_n = \frac{1}{2}\sum_{k=0}^{n} \frac{(-1)^k 2^k (n-k+2)(n-k+1)}{k!}\,.$$

SF13.20 Résoudre une équation différentielle

Il s'agit d'une utilisation importante des séries entières, que nous reverrons dans le thème 17 qui est consacré aux équations différentielles. Les calculs et les raisonnements sont semblables à ceux déjà faits en **[SF13.9]**, **[SF13.14]** et **[SF13.19]**).

Exemple

Déterminer une solution développable en série entière de l'équation différentielle :

$$x^2y'' + 4xy' + 2y = \ln(1+x)\,.$$

Soit f une solution développable en série entière au voisinage de 0, de rayon de convergence $R > 0$. On a donc, pour $x \in]-R\,;R[$:

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n \quad f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \quad \text{et} \quad f''(x) = \sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2}\,.$$

On dispose de plus du développement en série entière :

$$\forall\, x \in\ ]-1\,;1[\ ,\ \ln(1+x) = \sum_{n=1}^{+\infty} (-1)^{n-1}\frac{x^n}{n}\,.$$

Notons $R_1 = \min(R,1)$. f sera solution de l'équation différentielle proposée si et seulement si, pour tout $x \in]-R_1\,;R_1[$:

$$\sum_{\cancel{n=2}\ n=1}^{+\infty} n(n-1)a_n x^n + 4\sum_{n=1}^{+\infty} n a_n x^n + 2\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=1}^{+\infty} (-1)^{n-1}\frac{x^n}{n}\,.$$

✓ *Vous noterez l'astuce désormais classique consistant à faire partir la première somme à $n = 1$, ce qui est possible car le terme d'indice $n = 1$ est nul ; cela permet de faire débuter les sommes à partir du même indice.*

L'égalité précédente équivaut alors à :

$$\sum_{n=1}^{+\infty} [n(n-1) + 4n + 2]a_n x^n + 2a_0 = \sum_{n=1}^{+\infty} (-1)^{n-1}\frac{x^n}{n}$$

d'où, par unicité du développement en série entière :

$$a_0 = 0 \quad \text{et} \quad \forall\, n \in \mathbb{N}^*\ ,\ (n+1)(n+2)a_n = \frac{(-1)^{n-1}}{n}\,.$$

Réciproquement, la série entière $\displaystyle\sum_{n\geqslant 1} (-1)^{n-1}\frac{x^n}{n(n+1)(n+2)}$ est de rayon de convergence 1 (par exemple par la règle de d'Alembert) ; elle définit donc une fonction f de classe $\mathscr{C}^\infty$ sur $]-1\,;1[$. Il suffit alors de remonter les calculs précédents pour montrer que cette fonction est bien solution.

Il reste à expliciter $f(x) = \sum_{n=1}^{+\infty} (-1)^{n-1} \dfrac{x^n}{n(n+1)(n+2)}$ *pour* $x \in]-1\,;1[$.
Pour cela, nous allons utiliser une des techniques de sommation exposées en ***[SF13.5]*** *(il y a d'autres façons de faire que celle que nous avons choisie, nous vous invitons à les chercher !)*
On essaye ici de retrouver directement des séries entières connues ; pour cela, on effectue une décomposition en éléments simples, c'est-à-dire que l'on cherche des réels a, b, c *tels que :*

$$\forall\, n \in \mathbb{N}^*\ ,\ \frac{1}{n(n+1)(n+2)} = \frac{a}{n} + \frac{b}{n+1} + \frac{c}{n+2}\cdot$$

Pour obtenir a, b *et* c, *il suffit de réduire au même dénominateur et d'identifier les coefficients ; on obtient un système qu'il est facile de résoudre et l'on trouve* $a = c = \dfrac{1}{2}$ *et* $b = -1$. *On a donc, pour tout* $x \in\ ]-1\,;1[$, $x \neq 0$ *:*

$$\begin{aligned}
f(x) &= \frac{1}{2}\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n} - \sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n+1} + \frac{1}{2}\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n+2}\\
&= \frac{\ln(1+x)}{2} - \sum_{n=2}^{+\infty}(-1)^{n}\frac{x^{n-1}}{n} + \frac{1}{2}\sum_{n=3}^{+\infty}(-1)^{n-1}\frac{x^{n-2}}{n}\\
&= \frac{\ln(1+x)}{2} + \frac{1}{x}\sum_{n=2}^{+\infty}(-1)^{n-1}\frac{x^n}{n} + \frac{1}{2x^2}\sum_{n=3}^{+\infty}(-1)^{n-1}\frac{x^n}{n}\\
&= \frac{\ln(1+x)}{2} + \frac{1}{x}\left(\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n} - x\right)\\
&\qquad + \frac{1}{2x^2}\left(\sum_{n=1}^{+\infty}(-1)^{n-1}\frac{x^n}{n} - x + \frac{x^2}{2}\right)\\
&= \frac{\ln(1+x)}{2} + \frac{1}{x}\Big(\ln(1+x) - x\Big) + \frac{1}{2x^2}\left(\ln(1+x) - x + \frac{x^2}{2}\right)\\
&= \frac{\ln(1+x)}{2} + \left(x + \frac{1}{2}\right)\frac{\ln(1+x) - x}{x^2} + \frac{1}{4}\cdot
\end{aligned}$$

Les calculs précédents ne sont valables que pour $x \neq 0$, *mais la fonction obtenue est en fait de classe* $\mathscr{C}^\infty$ *sur* $]-1\,;1[$ *puisque* $x \mapsto \dfrac{\ln(1+x) - x}{x^2}$ *se prolonge en une fonction de classe* $\mathscr{C}^\infty$ *sur* $]-1\,;1[$ *car développable en série entière (il suffit, pour s'en convaincre, d'écrire le développement en série entière de* $x \mapsto \ln(1+x)$*).*

✓ *Les calculs ci-dessus sont donnés sans plus d'explication, nous vous invitons à les étudier soigneusement pour bien comprendre les manipulations qui ont été faites (changements d'indices).*

À vous de jouer !

Exercice 13.1 **SF11.10** **SF13.4**

Soit $a_n = \sin\left(\dfrac{1}{(-1)^n + \sqrt{n}}\right)$ $(n \geqslant 2)$. Quel est le domaine de définition de $\displaystyle\sum_{n=2}^{+\infty} a_n x^n$?

Exercice 13.2 **SF13.3** **SF13.6**

Déterminer le rayon de convergence et somme de la série entière $\displaystyle\sum_{n=1}^{+\infty} n^{(-1)^n} x^n$.

Exercice 13.3 **SF13.1** **SF13.3** **SF13.6**

Soit $\displaystyle\sum_{n\in\mathbb{N}} a_n z^n$ une série entière de rayon de convergence $R > 0$.
Déterminer les rayons de convergence des séries entières :

1) $\displaystyle\sum_{n\in\mathbb{N}} a_n^2 z^n$ **2)** $\displaystyle\sum_{n\in\mathbb{N}} \frac{a_n}{n!} z^n$.

Exercice 13.4 **SF13.1** **SF13.2** **SF13.3**

Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres réels strictement positifs, telle que la série $\sum a_n$ diverge. Pour tout $n \in \mathbb{N}^*$, on note $s_n = \displaystyle\sum_{k=0}^{n} a_k$, et on suppose que $\displaystyle\lim_{n\to+\infty} \frac{a_n}{s_n} = 0$.
Montrer que les séries entières $\displaystyle\sum_{n\geqslant 0} a_n z^n$ et $\displaystyle\sum_{n\geqslant 0} s_n z^n$ ont même rayon de convergence, que l'on calculera, et comparer leurs sommes.

Exercice 13.5 **SF13.14**

Montrer que la fonction $f\colon x \mapsto \dfrac{\operatorname{Arcsin} x}{\sqrt{1-x^2}}$ est développable en série entière sur $]-1\,;1[$ et calculer ce développement.

Exercice 13.6 **SF12.20** **SF13.2** **SF13.9**

1) Calculer $I_{n,m} = \displaystyle\int_0^1 t^n(1-t)^m \,\mathrm{d}t$ pour $(n,m) \in \mathbb{N}^2$.

Pour $n \in \mathbb{N}$, on pose $a_n = I_{n,n} = \displaystyle\int_0^1 t^n(1-t)^n \,\mathrm{d}t$.

2) Calculer le rayon de convergence R de la série entière $\displaystyle\sum_{n\geqslant 0} a_n x^n$.

3) Calculer la somme de cette série quand elle existe.

Exercice 13.7 — SF13.18

Calculer l'intégrale : $I = \displaystyle\int_0^1 \ln t \ln(1-t)\,\mathrm{d}t$.

Exercice 13.8 — SF13.2 SF13.9

Déterminer le rayon de convergence et calculer la somme de la série entière :

$$\sum_{n\geqslant 0} \frac{4^n (n!)^2}{(2n+1)!} x^{2n+1} .$$

Exercice 13.9 — SF12.11 SF12.13 SF13.4

Soit $f : x \mapsto \displaystyle\sum_{n=1}^{+\infty} \sin\left(\frac{1}{\sqrt{n}}\right) x^n$.

1) Déterminer le domaine de définition de f.

2) Déterminer la limite de $f(x)$ quand $x \to 1^-$.

3) Montrer que $\displaystyle\lim_{x\to 1^-} (1-x) f(x) = 0$.

Exercice 13.10 — SF13.5 SF13.19

On définit une suite réelle (a_n) en posant : $a_0 = 1$, $2a_{n+1} = \displaystyle\sum_{k=0}^{n} \binom{n}{k} a_k a_{n-k}$.

En considérant une série entière convenable, calculer a_n en fonction de n.

Thème 14 - Intégration sur un intervalle quelconque

■ Comment étudier une intégrale généralisée ?

- Pour exposer les méthodes d'étude, nous nous placerons dans le cadre suivant : f est une fonction définie sur un intervalle $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs dans $\mathbb{R}$.

  Vous adapterez sans difficulté ces méthodes au cas d'une fonction définie sur un intervalle $]a\,;b]$ avec $-\infty \leqslant a < b < +\infty$. Dans le cas d'une fonction définie sur un intervalle ouvert $]a\,;b[$, rappelons que $\displaystyle\int_{]a;b[} f$ est dite convergente si et seulement si pour tout/un $c \in\ ]a\,;b[$, les deux intégrales $\displaystyle\int_{]a;c]} f$ et $\displaystyle\int_{[c;b[} f$ le sont ; le problème se ramène donc alors aux précédents.
- **La première chose à faire** est de vérifier que f est bien continue par morceaux sur $[a\,;b[$, en particulier que $f(t)$ est bien défini pour tout $t \in [a\,;b[$. Ainsi, pour tout $c \in [a\,;b[$, l'intégrale $\displaystyle\int_a^c f$ existe et, pour la suite, on peut se contenter d'examiner le comportement de f au voisinage immédiat de b.

  ✓ Ce point est fondamental et trop souvent oublié ; non seulement il est indispensable afin de pouvoir appliquer les théorèmes du cours, mais il permet aussi de détecter les problèmes, en particulier lorsqu'il y a un paramètre.

  ✓ Dans le cas où f est une application continue par morceaux sur le segment $[a\,;b]$ avec a et b finis, l'intégrale $\displaystyle\int_a^b$ de f existe et ce n'est pas une intégrale impropre. Aucune étude spéciale n'est donc nécessaire.

SF14.1 Reconnaître une intégrale faussement impropre

Dans le cas où l'intervalle $[a\,;b[$ est borné (c'est-à-dire b fini), on peut parfois utiliser le résultat suivant.

– Si f admet une limite finie en b^-, elle est intégrable.
– Plus généralement, si f est bornée sur $[a\,;b[$, elle y est intégrable.

✓ Bien sûr ce critère n'est pas applicable lorsque l'intervalle n'est pas borné. Par exemple, l'intégrale $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ diverge bien que $\displaystyle\lim_{t\to+\infty} \frac{1}{t} = 0$.

✓ Ce critère ne fournit qu'une condition *suffisante* d'existence ; par exemple, l'intégrale $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}$ converge bien que $\displaystyle\lim_{t\to0^+} \frac{1}{\sqrt{t}} = +\infty$.

Exemple

Justifier l'existence de $\displaystyle\int_0^1 \frac{t\ln t}{1-t}\,\mathrm{d}t$.

- *On commence par remarquer que, d'après les théorèmes usuels, la fonction* $f : t \mapsto \dfrac{t \ln t}{1-t}$ *est continue sur* $]0\,;1[$.
- *D'après les résultats sur les croissances comparées des fonctions usuelles,* $\lim\limits_{t\to 0^+} f(t) = 0$. *On connaît aussi la limite :* $\lim\limits_{t\to 1} \dfrac{\ln t}{t-1} = 1$, *donc* $\lim\limits_{t\to 1^-} f(t) = -1$.
- *Ainsi* f *se prolonge en une fonction continue sur* $[0\,;1]$ *et y est intégrable.*

Exemple

Justifier l'existence de $\displaystyle\int_0^1 \left(\frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor\right) \mathrm{d}t$.

- *La fonction « partie entière » étant continue par morceaux sur* $\mathbb{R}$, *la fonction* $f : t \mapsto \dfrac{1}{t} - \left\lfloor \dfrac{1}{t} \right\rfloor$ *est continue par morceaux sur* $]0\,;1]$ *(plus précisément elle est continue sur chaque intervalle* $\left]\dfrac{1}{n+1}\,;\dfrac{1}{n}\right]$ *avec* $n \in \mathbb{N}^*$*).*
- *Puisque pour tout* x *réel,* $0 \leqslant x - \lfloor x \rfloor < 1$, f *est bornée (comprise entre* 0 *et* 1*) ; elle est donc intégrable sur* $]0\,;1]$.

  ✓ *Cependant,* f *n'admet pas de limite en* 0^+ *: pour le démontrer, considérer les suites de terme général* $\dfrac{1}{n}$ *et* $\dfrac{1}{n+1/2}$, *et utiliser le théorème sur la caractérisation séquentielle de la limite.*

SF14.2 Revenir à la définition

Dans le cas où l'on connaît une primitive F de f, il suffit d'examiner si $\displaystyle\int_a^x f(t)\,\mathrm{d}t = F(x) - F(a)$ possède une limite (finie) en b^-.

✓ Ce cas est assez peu fréquent ; il faut y penser lorsque l'on vous demande non seulement de prouver l'existence d'une intégrale généralisée mais aussi de la calculer. Mais même dans ce cas nous conseillons d'étudier d'abord l'existence afin d'être certain que les calculs qui suivent sont valides.

Exemple

Montrer que l'intégrale $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^2}}$ existe et calculer sa valeur.

La fonction $t \mapsto \dfrac{1}{\sqrt{1-t^2}}$ *est définie et continue sur* $[0\,;1[$, *par les théorèmes usuels.*
Pour tout $x \in [0\,;1[$ *on a :* $\displaystyle\int_0^x \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \big[\operatorname{Arcsin} t\big]_0^1 = \operatorname{Arcsin} x$,
d'où l'on déduit l'existence de l'intégrale proposée et sa valeur :

$$\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \lim_{x\to 1^-} \operatorname{Arcsin} x = \operatorname{Arcsin} 1 = \frac{\pi}{2}\cdot$$

SF14.3 Utiliser les théorèmes de comparaison

L'idée est ici de comparer la fonction à intégrer à une fonction plus simple. Il s'agit de la méthode la plus utile si l'on excepte les cas simples vus auparavant.

- Les théorèmes de comparaison rappelés en **[S14.3]** concernent les fonctions *à valeurs réelles positives* (au moins au voisinage de b^-).
  Si ce n'est pas le cas (ou si l'on ne sait pas déterminer le signe de f, ou si f est à valeurs complexes), on appliquera ces théorèmes à $|f|$ (valeur absolue ou module) puisque l'on sait qu'une intégrale absolument convergente est convergente.

- Cependant, l'étude du signe de f ne doit pas être systématique : cela dépend du critère employé. En effet :
  - si $f \underset{t\to b^-}{\sim} \lambda\, g$, où g est une fonction positive au voisinage de b et λ un réel non nul, le signe de f sera celui de λ au voisinage de b ;
  - si $f \underset{t\to b^-}{=} o(g)$ ou $f \underset{t\to b^-}{=} O(g)$, où g est *positive* et intégrable au voisinage de b, on a en fait, en revenant à la définition, $|f(t)| \leqslant kg(t)$ au voisinage de b avec k constante, et l'intégrale de f sera directement absolument convergente.

- Parmi les méthodes de comparaison, l'utilisation d'équivalents est la méthode à laquelle il faut penser en premier (ne serait-ce que mentalement). Ce n'est que lorsqu'elle a échoué que l'on peut penser à utiliser les « $o(\ )$ » ou les « $O(\ )$ », et ensuite une inégalité, qui est, sauf exception, plus délicate à obtenir.

- La comparaison la plus fréquemment utilisée est celle avec les intégrales de Riemann $\displaystyle\int_0^a \frac{\mathrm{d}t}{t^\alpha}$ ou $\displaystyle\int_a^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ ou $\displaystyle\int_a^b \frac{\mathrm{d}t}{(b-t)^\alpha}$; elle consiste à étudier la limite de $t^\alpha f(t)$ lorsque t tend vers 0 ou $\pm\infty$, ou celle de $(b-t)^\alpha f(t)$ lorsque t tend vers b^-. Cette méthode rappelle celle concernant les séries **[SF11.5]** et s'utilise de préférence lorsqu'apparaît dans l'expression de f des exponentielles ou des logarithmes, qu'il est facile de comparer à t^α.

- Très souvent, il faut savoir mêler plusieurs des méthodes ci-dessus : par exemple, on peut d'abord chercher un équivalent, puis appliquer l'une des autres méthodes à la fonction plus simple ainsi obtenue.

Exemple

Étudier l'existence de $\displaystyle\int_1^{+\infty} \frac{\mathrm{e}^{-x}}{\sqrt{x^2-1}}\,\mathrm{d}x$.

Soit $f : x \mapsto \dfrac{\mathrm{e}^{-x}}{\sqrt{x^2-1}}$ *pour* $x > 1$.

• *On commence* déjà *par remarquer que* f *est définie et continue sur* $]1\,;+\infty[$ *: cela résulte des théorèmes usuels, nous ne détaillons pas.*

✓ *Commencer l'étude d'une intégrale impropre en se précipitant sur l'étude aux bornes sans avoir auparavant étudié le domaine de définition et la continuité de la fonction à intégrer est une faute grave !*

• f *étant continue sur* $]1\,;+\infty[$*, les problèmes d'intégrabilité ne se posent donc qu'au voisinage de* 1^+ *et de* $+\infty$.

– Au voisinage de 1^+

On commence par examiner la fonction et se « débarrasser » des parties qui ne posent pas de problème. Ici, $\lim\limits_{x\to 1} \mathrm{e}^{-x} = \dfrac{1}{\mathrm{e}}$ *donc on a* $f(x) \underset{x\to 1^+}{\sim} \dfrac{\mathrm{e}^{-1}}{\sqrt{x^2-1}}$. *Ensuite, on remarque que :*

$$\sqrt{x^2-1} = \sqrt{(x-1)(x+1)} = \sqrt{x-1}\sqrt{x+1}\,.$$

Là encore, l'expression $\sqrt{x+1}$ *ne pose pas de problème au voisinage de* 1 *puisque* $\lim\limits_{x\to 1}\sqrt{x+1} = \sqrt{2}$, *et on a finalement :* $f(x) \underset{x\to 1^+}{\sim} \dfrac{\mathrm{e}^{-1}}{\sqrt{2}\sqrt{x-1}}$.

Or, la fonction $x \mapsto \dfrac{1}{\sqrt{x-1}} = \dfrac{1}{(x-1)^{1/2}}$ *est intégrable au voisinage de* 1^+ *:*

c'est une intégrale de Riemann, de même nature que celle de $x \mapsto \dfrac{1}{x^{1/2}}$ *au voisinage de* 0^+, *donc convergente* ($1/2 < 1$).
Puisqu'il s'agit d'une fonction positive (critère à ne pas oublier !), les théorèmes de comparaison (ici, l'utilisation de fonctions équivalentes) assurent que f *est aussi intégrable au voisinage de* 1^+.

– Au voisinage de $+\infty$

Là encore on examine la fonction, pour remarquer que le terme le plus important ici est e^{-x}. *Cela suggère d'utiliser une comparaison avec une intégrale de Riemann.*

Plus précisément, $f(x) \underset{x\to+\infty}{\sim} \dfrac{\mathrm{e}^{-x}}{x}$ *donc, d'après les croissances comparées,* $\lim\limits_{x\to+\infty} x^2 f(x) = 0$.

Donc $f(x) \underset{x\to+\infty}{=} o\left(\dfrac{1}{x^2}\right)$ *; puisque* $x \mapsto \dfrac{1}{x^2}$ *est positive et intégrable au voisinage de* $+\infty$, *il en est de même de* f.

En rassemblant les deux cas, on conclut : f *est intégrable sur* $]1\,;+\infty[$.

✓ Nous avons volontairement très détaillé la réponse à cet exemple, afin que vous compreniez la démarche à adopter. En particulier, il est important de distinguer dans une fonction quels sont les termes qui posent problème, et les autres que l'on peut remplacer par un équivalent.

Exemple

Soit a un réel de l'intervalle $]0\,;1[$. Étudier selon les valeurs des réels α et β l'existence de l'intégrale : $\displaystyle\int_0^a \frac{\mathrm{d}t}{t^\alpha\,|\ln t|^\beta}$.

L'intégrale proposée s'appelle une intégrale de Bertrand. *Les résultats concernant ces intégrales ne figurent pas au programme mais la méthode d'étude (comparaison à une intégrale de Riemann) est très instructive, s'applique à de nombreuses autres fonctions, et mérite que vous l'étudiez de façon approfondie.*

- *Tout d'abord, on note que la fonction* $t \mapsto \dfrac{1}{t^\alpha\,|\ln t|^\beta}$ *est continue sur* $]0\,;a[$, *et qu'elle est à valeurs positives.*
- *Ensuite, on remarque que, dans le cas* $\alpha = 1$, *on peut expliciter une primitive.*
  – *Si* $\beta = 1$, $\displaystyle\int_b^a \frac{\mathrm{d}t}{t\,|\ln t|} = -\int_b^a \frac{\mathrm{d}t}{t\ln t} = \big[-\ln|\ln t|\big]_b^a = \ln|\ln b| - \ln|\ln a|$,
  donc $\lim\limits_{b\to 0^+}\displaystyle\int_b^a \frac{\mathrm{d}t}{t\,|\ln t|} = +\infty$ *: l'intégrale diverge.*

– *Si* $\beta \neq 1$,

$$\int_b^a \frac{\mathrm{d}t}{t\,|\ln t|^\beta} = \int_b^a \frac{\mathrm{d}t}{t(-\ln t)^\beta} = \left[-\frac{|\ln t|^{-\beta+1}}{-\beta+1}\right]_b^a = \frac{|\ln b|^{-\beta+1} - |\ln a|^{-\beta+1}}{-\beta+1}$$

et, puisque $\lim\limits_{b\to 0^+} |\ln b| = +\infty$, *la limite de l'intégrale précédente quand* $b \to 0^+$ *est* $+\infty$ *si* $-\beta+1 > 0$ *et* $-\dfrac{|\ln a|^{-\beta+1}}{-\beta+1}$ *si* $-\beta+1 < 0$.

Donc l'intégrale converge dans le cas $\alpha = 1$ *si et seulement si* $\beta > 1$.

- *Il reste à traiter les cas* $\alpha \neq 1$ *; la méthode consiste à comparer à une intégrale de Riemann* $\displaystyle\int_0^a \frac{\mathrm{d}t}{t^\gamma}$, *où* γ *est un réel bien choisi.*

  – *Si* $\alpha < 1$, *on choisit* γ *tel que* $\alpha < \gamma < 1$.

  Alors $t^\gamma \dfrac{1}{t^\alpha\,|\ln t|^\beta} = t^{\gamma-\alpha}\,|\ln t|^{-\beta}$ *et, puisque* $\gamma - \alpha > 0$, *les résultats sur les croissances comparées des fonctions usuelles donnent* $\lim\limits_{t\to 0^+} t^\gamma \dfrac{1}{t^\alpha\,|\ln t|^\beta} = 0$.

  Cela signifie que $\dfrac{1}{t^\alpha\,|\ln t|^\beta}$ *est négligeable au voisinage de* 0^+ *devant* $\dfrac{1}{t^\gamma}$.

  Or, puisque $\gamma < 1$, *la fonction (positive)* $t \mapsto \dfrac{1}{t^\gamma}$ *est intégrable au voisinage de* 0^+, *et il en est donc de même de* $t \mapsto \dfrac{1}{t^\alpha\,|\ln t|^\beta}$.

  – *Si* $\alpha > 1$, *on choisit* $\gamma > 1$ *tel que* $1 < \gamma < \alpha$.

  On aura maintenant $\lim\limits_{t\to 0^+} t^\gamma \dfrac{1}{t^\alpha\,|\ln t|^\beta} = +\infty$, *ce qui implique qu'il existe un voisinage de* 0^+ *sur lequel* $t^\gamma \dfrac{1}{t^\alpha\,|\ln t|^\beta} > 1$, *soit* $\dfrac{1}{t^\alpha\,|\ln t|^\beta} > \dfrac{1}{t^\gamma}$. *Or, puisque* $\gamma > 1$, *la fonction* $t \mapsto \dfrac{1}{t^\gamma}$ *n'est pas intégrable au voisinage de* 0^+, *et, par comparaison des intégrales de fonctions positives, il en est de même de* $t \mapsto \dfrac{1}{t^\alpha\,|\ln t|^\beta}$.

En conclusion, si $0 < a < 1$ *:*

$$\int_0^a \frac{\mathrm{d}t}{t^\alpha\,|\ln t|^\beta} \text{ converge} \iff \alpha < 1 \text{ ou } (\alpha = 1 \text{ et } \beta > 1)\,.$$

- *Nous vous laissons le soin de démontrer, à titre d'entraînement pour bien maîtriser la méthode, que, si* $a > 1$ *:*

$$\int_a^{+\infty} \frac{\mathrm{d}t}{t^\alpha(\ln t)^\beta} \text{ converge} \iff \alpha > 1 \text{ ou } (\alpha = 1 \text{ et } \beta > 1)\,.$$

La méthode à suivre est similaire.

– *Dans le cas* $\alpha = 1$, *on peut calculer une primitive.*

– *Dans le cas* $\alpha \neq 1$ *on étudie* $\lim\limits_{t\to+\infty} t^\gamma \dfrac{1}{t^\alpha(\ln t)^\beta}$ *où* γ *est un réel compris strictement entre* α *et* 1 *; selon les cas, cette limite est* 0 *ou* $+\infty$, *et on peut alors conclure en comparant avec l'intégrale de Riemann* $\displaystyle\int_a^{+\infty} \frac{\mathrm{d}t}{t^\gamma}$.

SF14.4 Utiliser une intégration par parties

Dans le cas d'une intégrale de la forme $\displaystyle\int_a^{+\infty} \frac{f(t)}{t^\alpha}\,\mathrm{d}t$ avec $\alpha > 0$, une intégration par parties en dérivant $t \mapsto \dfrac{1}{t^\alpha}$ fait apparaître un terme en $\dfrac{1}{t^{\alpha+1}}$, avec ici $\alpha+1 > 1$, ce qui, par comparaison à une intégrale de Riemann, peut permettre de conclure.

Exemple

Soit a, b deux réels avec $a \neq 0$ et α un réel strictement positif.

Montrer que l'intégrale $\displaystyle\int_1^{+\infty} \frac{\sin(at+b)}{t^\alpha}\,\mathrm{d}t$ est convergente.

On note déjà que la fonction $t \mapsto \dfrac{\sin(at+b)}{t^\alpha}$ *est continue sur* $[1\,;+\infty[$.
Soit donc x *un réel* > 1. *Une intégration par parties donne :*

$$\int_1^x \frac{\sin(at+b)}{t^\alpha}\,\mathrm{d}t = \left[-\frac{1}{a}\frac{\cos(at+b)}{t^\alpha}\right]_1^x - \frac{\alpha}{a}\int_1^x \frac{\cos(at+b)}{t^{\alpha+1}}\,\mathrm{d}t \quad (*)$$

Or l'intégrale $\displaystyle\int_1^{+\infty} \frac{\cos(at+b)}{t^{\alpha+1}}\,\mathrm{d}t$ *est absolument convergente, puisque* $\left|\dfrac{\cos(at+b)}{t^{\alpha+1}}\right| \leqslant \dfrac{1}{t^{\alpha+1}}$ *avec* $\alpha+1 > 1$ *; elle est donc convergente, c'est-à-dire que* $\displaystyle\lim_{x\to+\infty}\int_1^x \frac{\cos(at+b)}{t^{\alpha+1}}\,\mathrm{d}t$ *existe.*

De plus $\displaystyle\lim_{x\to+\infty} \frac{\cos(ax+b)}{x^\alpha} = 0$ *(car* $\alpha > 0$ *et la fonction* cos *est bornée).*

D'après ce qui précède et la relation $(*)$, $\displaystyle\lim_{x\to+\infty}\int_1^x \frac{\sin(at+b)}{t^\alpha}\,\mathrm{d}t$ *existe, ce qu'il fallait démontrer.*

SF14.5 Utiliser un développement limité

Cette méthode consiste à écrire $f(t) \underset{t\to b^-}{=} g(t) + h(t)$ à l'aide d'un développement limité ; si les deux intégrales de g et de h convergent, il en sera de même de celle de f ; si l'une des intégrales est convergente et l'autre divergente, l'intégrale de f sera divergente.

Exemple

Nature de l'intégrale : $\displaystyle\int_0^{+\infty} \frac{\sin t}{\sqrt{t}+\cos t}\,\mathrm{d}t$.

Soit $f\colon t \mapsto \dfrac{\sin t}{\sqrt{t}+\cos t}$.

- *Pour démontrer la continuité de* f *sur* $\mathbb{R}_+$, *il suffit de prouver que l'expression* $\sqrt{t}+\cos t$ *ne s'annule pas.*
  Or, pour $t \in [0\,;1]$, $\sqrt{t} \geqslant 0$ *et* $\cos t > 0$ *donc* $\sqrt{t}+\cos t > 0$ *et pour* $t > 1$, $\sqrt{t} > 1$ *donc on a encore* $\sqrt{t}+\cos t > 0$.
  f *est donc bien définie et continue sur* $\mathbb{R}_+$, *et le problème de l'intégrabilité ne se pose qu'au voisinage de* $+\infty$.
- *Avec un peu de réflexion, on s'aperçoit qu'aucune des méthodes précédemment évoquées ne s'applique ici.*

On effectue donc un développement limité de f au voisinage de $+\infty$:

$$\frac{\sin t}{\sqrt{t}+\cos t} = \frac{\sin t}{\sqrt{t}\left(1+\frac{\cos t}{\sqrt{t}}\right)} = \frac{\sin t}{\sqrt{t}}\left(1-\frac{\cos t}{\sqrt{t}}+O\left(\frac{1}{t}\right)\right)$$
$$= \frac{\sin t}{\sqrt{t}} - \frac{1}{2}\frac{\sin 2t}{t} + O\left(\frac{1}{t^{3/2}}\right).$$

D'après l'exemple traité ci-dessus en **[SF14.4]**, *les intégrales* $\displaystyle\int_1^{+\infty} \frac{\sin t}{\sqrt{t}}\,\mathrm{d}t$ *et* $\displaystyle\int_1^{+\infty} \frac{\sin 2t}{t}\,\mathrm{d}t$ *sont convergentes ; d'après les théorèmes de comparaison, le terme en* $O\left(\frac{1}{t^{3/2}}\right)$ *est intégrable au voisinage de* $+\infty$.

On en déduit que $\displaystyle\int_1^{+\infty} \frac{\sin t}{\sqrt{t}+\cos t}\,\mathrm{d}t$ *existe (somme de trois intégrales convergentes), donc l'intégrale proposée aussi.*

SF14.6 Utiliser un changement de variable

Pour faire un changement de variable dans une intégrale impropre, on dispose du résultat rappelé en **[S14.5]**. Cependant, lorsqu'on ne sait pas que les intégrales considérées existent, il est plus sûr de faire le changement de variable dans l'intégrale sur un segment, puis de passer à la limite lorsque le calcul est terminé.

Exemple

Étudier l'existence de $\displaystyle\int_0^{+\infty} \sin t^2\,\mathrm{d}t$.

La fonction $t \mapsto \sin^2 t$ *est continue sur* $\mathbb{R}_+$ *; pour tout* $x \geqslant 0$,

$$\int_0^x \sin t^2\,\mathrm{d}t = \int_0^{x^2} \frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u,$$

en ayant posé $u = t^2$.

Or on sait que l'intégrale $\displaystyle\int_0^{+\infty} \frac{\sin u}{\sqrt{u}}\,\mathrm{d}u$ *est convergente (voir exemple page 541), donc* $\displaystyle\lim_{x\to+\infty} \int_0^{x^2} \frac{\sin u}{2\sqrt{u}}\,\mathrm{d}u$ *existe, et l'intégrale proposée est convergente par définition.*

SF14.7 Utiliser les séries

Il existe principalement deux résultats pour relier séries et intégrales.

- Nous avons déjà vu en **[S11.7]** la méthode de comparaison série/intégrale.
- Un autre résultat peut être utile :

  Soit f une fonction définie sur un intervalle de la forme $[a\,;b[$, avec $-\infty < a < b \leqslant +\infty$, à valeurs réelles *positives*.
  Soit $(x_n)_{n\in\mathbb{N}}$ une suite croissante d'éléments de $[a\,;b[$ telle que $\displaystyle\lim_{n\to+\infty} x_n = b$.

  Alors l'intégrale $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ converge si et seulement si la série de terme général $u_n = \displaystyle\int_{x_n}^{x_{n+1}} f(t)\,\mathrm{d}t$ converge.

Ce résultat est intuitivement facile à comprendre, mais il ne figure pas au programme ; il faut donc le redémontrer dans chaque cas. C'est ce que nous faisons dans l'exemple suivant.

Exemple

Justifier l'existence de l'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{1+t^4\sin^2 t}$.

La fonction $t \mapsto \dfrac{1}{1+t^4\sin^2 t}$ *est continue sur* $\mathbb{R}_+$. *Pour tout* $x \geqslant 0$, *posons :* $F(x) = \displaystyle\int_0^x \frac{\mathrm{d}t}{1+t^4\sin^2 t}$. *Il s'agit donc d'étudier* $\lim\limits_{x\to+\infty} F(x)$.

- *Soit* $n \in \mathbb{N}^*$. *Par la relation de Chasles,* $F(n\pi) = \displaystyle\sum_{k=0}^{n-1} u_k$ *où* $u_k = \displaystyle\int_{k\pi}^{(k+1)\pi} \frac{\mathrm{d}t}{1+t^4\sin^2 t}$.
  En faisant le changement de variable $u = t - k\pi$, *on obtient :*
  $$u_k = \int_0^{\pi} \frac{\mathrm{d}u}{1+(u+k\pi)^4\sin^2 u}$$
  d'où l'encadrement :
  $$\int_0^{\pi} \frac{\mathrm{d}u}{1+((k+1)\pi)^4\sin^2 u} \leqslant u_k \leqslant \int_0^{\pi} \frac{\mathrm{d}u}{1+(k\pi)^4\sin^2 u} \quad (*)$$
  Or pour a *paramètre réel positif, on peut calculer :*
  $$\int_0^{\pi} \frac{\mathrm{d}u}{1+a\sin^2 u} = \int_0^{\pi} \frac{1}{\frac{1}{\sin^2 u}+a} \cdot \frac{\mathrm{d}u}{\sin^2 u} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}v}{v^2+a+1}$$
  $$= \frac{1}{\sqrt{a+1}} \left[\operatorname{Arctan} \frac{v}{\sqrt{a+1}}\right]_{-\infty}^{+\infty} = \frac{\pi}{\sqrt{a+1}}$$
  (on a fait le changement de variable $v = \cot u$*).*
  En remplaçant dans $(*)$ *on obtient l'encadrement*
  $$\frac{\pi}{\sqrt{((k+1)\pi)^4+1}} \leqslant u_k \leqslant \frac{\pi}{\sqrt{(k\pi)^4+1}}$$
  qui permet de conclure : $u_k \underset{k\to+\infty}{=} O\left(\dfrac{1}{k^2}\right)$.
  D'après les théorèmes de comparaison des séries à termes positifs, la série $\sum u_k$ *est convergente, ce qui signifie que* $\lim\limits_{n\to+\infty} F(n\pi)$ *existe.*
- *Notons* $\ell = \lim\limits_{n\to+\infty} F(n\pi)$. *Pour tout* $x \geqslant 0$, *il existe un et un seul entier* n_x *tel que* $n_x\pi \leqslant x < (n_x+1)\pi$ *; plus précisément,* $n_x = \left\lfloor \dfrac{x}{\pi} \right\rfloor$.
  La fonction F *étant croissante (car c'est une primitive d'une fonction positive), on aura l'encadrement :*
  $$F(n_x\pi) \leqslant F(x) \leqslant F((n_x+1)\pi)$$
  et puisque n_x *tend vers* $+\infty$ *avec* x, *on en déduit* $\lim\limits_{x\to+\infty} F(x) = \ell$.

En conclusion, l'intégrale proposée existe.

■ Comment étudier l'intégrabilité dans un cadre plus théorique ?

Nous avons déjà abordé cette problématique dans un chapitre précédent page 444, dans le cadre des intégrales sur un segment, et les méthodes exposées restent valables (intégration par parties, changement de variables,...). Nous donnons ici quelques pistes supplémentaires lorsqu'il s'agit d'intégrales généralisées.

SF14.8 Utiliser la définition

Inutile parfois de chercher des choses compliquées ! Dire que l'intégrale de f sur $[a\,;b[$ existe équivaut à dire que $\lim\limits_{x\to b^-}\int_a^x f(t)\,\mathrm{d}t$ existe, tout simplement.

Exemple

Soit f de classe $\mathscr{C}^2$ de $\mathbb{R}_+$ dans $\mathbb{R}$, telles que f et f'' soient intégrables sur $\mathbb{R}_+$.
Montrer que $\lim\limits_{x\to+\infty} f'(x)=0$, puis que $\lim\limits_{x\to+\infty} f(x)=0$.

- *Pour tout $x\in\mathbb{R}$ on a :* $f'(x)=f'(0)+\displaystyle\int_0^x f''(t)\,\mathrm{d}t$.
  f'' étant intégrable sur $\mathbb{R}_+$, $\lim\limits_{x\to+\infty}\int_0^x f''$ existe et par suite $\ell=\lim\limits_{x\to+\infty} f'(x)$ existe.
  Si par l'absurde cette limite était non nulle, par exemple $\ell>0$, il existerait $A\geqslant 0$ tel que $f'(x)\geqslant\dfrac{\ell}{2}$ pour tout $x\geqslant A$ et alors :
  $$\forall\, x\geqslant A,\ f(x)=f(A)+\int_A^x f'(t)\,\mathrm{d}t\geqslant f(A)+\frac{\ell}{2}(x-A)\,,$$
  d'où $\lim\limits_{x\to+\infty} f(x)=+\infty$: contradiction.
  Si $\ell<0$ on obtient aussi une contradiction. En conclusion, $\ell=0$.
- *L'inégalité de Taylor-Lagrange appliquée à une primitive F de f entre x et $x+1$ donne, pour tout $x\in\mathbb{R}_+$:* $\left|\displaystyle\int_x^{x+1} f-f(x)\right|\leqslant\frac{1}{2}\max_{t\in[x;x+1]}|f'(t)|$.
  Puisque $\displaystyle\int_0^{+\infty} f$ existe, on a $\lim\limits_{x\to+\infty}\int_x^{x+1} f=0$, et puisque $\lim\limits_{t\to+\infty} f'(t)=0$, on déduit de l'inégalité ci-dessus : $\lim\limits_{x\to+\infty} f(x)=0$.

Une autre conséquence de la définition, souvent utilisée, est que, si $\displaystyle\int_a^b f$ existe, alors $\lim\limits_{x\to b^-}\int_x^b f(t)\,\mathrm{d}t=0$. Nous nous en sommes d'ailleurs servi dans l'exemple précédent.

Exemple

Soit $f\colon\mathbb{R}_+\to\mathbb{R}$, uniformément continue et telle que $\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t$ existe.
Montrer que $\lim\limits_{x\to+\infty} f(x)=0$.

Soit $\varepsilon > 0$. Par définition de la continuité uniforme, il existe $\alpha > 0$ tel que :

$$\forall\, x,\, y \in \mathbb{R}_+,\ |x-y| \leqslant \alpha \implies |f(x)-f(y)| \leqslant \varepsilon\,.$$

On a donc, pour tout $x \geqslant 0$:

$$\left|\int_x^{x+\alpha} f(t)\,\mathrm{d}t - \alpha f(x)\right| = \left|\int_x^{x+\alpha} \big(f(t)-f(x)\big)\,\mathrm{d}t\right| \leqslant \int_x^{x+\alpha} |f(t)-f(x)|\,\mathrm{d}t \leqslant \alpha\varepsilon\,,$$

d'où l'on déduit, par l'inégalité triangulaire :

$$|\alpha f(x)| \leqslant \alpha\varepsilon + \left|\int_x^{x+\alpha} f(t)\,\mathrm{d}t\right|.$$

Or $\lim\limits_{x\to+\infty} \int_x^{x+\alpha} f(t)\,\mathrm{d}t = 0$ puisque l'intégrale de f sur $\mathbb{R}_+$ converge. Il existe donc A tel que pour $x > A$ on ait $\left|\int_x^{x+\alpha} f(t)\,\mathrm{d}t\right| \leqslant \alpha\varepsilon$ et finalement :

$$\forall\, x \geqslant A,\ |f(x)| \leqslant 2\varepsilon\,.$$

Cela prouve que $\lim\limits_{x\to+\infty} f(x) = 0$.

SF14.9 Utiliser une primitive

On exploite encore l'idée précédente. Si la limite $\lim\limits_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t$ existe, la fonction $x \mapsto \int_a^x f(t)\,\mathrm{d}t$, qui est une primitive de f, est donc bornée sur $[a\,;b[$.

Exemple

Soit f une fonction continue sur $\mathbb{R}_+$ à valeurs dans $\mathbb{R}$.

On suppose qu'il existe $a \in \mathbb{R}$ tel que $\int_0^{+\infty} f(t)\mathrm{e}^{-at}\,\mathrm{d}t$ existe.

Montrer que pour tout $b \geqslant a$, l'intégrale $\int_0^{+\infty} f(t)\mathrm{e}^{-bt}\,\mathrm{d}t$ existe.

- *Le résultat serait immédiat en remplaçant l'hypothèse « l'intégrale existe » par : « est intégrable », puisque pour $b \geqslant a$, $|f(t)|\,\mathrm{e}^{-bt} \leqslant |f(t)|\,\mathrm{e}^{-at}$!*

  ✓ *Nous avons fait cette remarque pour vous encourager à être attentif dans la lecture de l'énoncé...*

- *Le cas $b = a$ étant évident, nous supposerons ici $b > a$.*
  Pour tout $x \in \mathbb{R}_+$, posons : $F(x) = \int_0^x f(t)\mathrm{e}^{-at}\,\mathrm{d}t$. F est de classe $\mathscr{C}^1$ et $F'(x) = f(x)\mathrm{e}^{-ax}$. Par intégration par parties :

$$\begin{aligned}\int_0^x f(t)\mathrm{e}^{-bt}\,\mathrm{d}t &= \int_0^x F'(t)\mathrm{e}^{(a-b)t}\,\mathrm{d}t\\ &= \big[F(t)\mathrm{e}^{(a-b)t}\big]_0^x + (b-a)\int_0^x F(t)\mathrm{e}^{(a-b)t}\,\mathrm{d}t\\ &= F(x)\mathrm{e}^{(a-b)x} + (b-a)\int_0^x F(t)\mathrm{e}^{(a-b)t}\,\mathrm{d}t\,.\end{aligned}$$

F étant bornée sur $\mathbb{R}_+$ (car continue et admet une limite en $+\infty$), $\lim\limits_{x\to+\infty} F(x)\mathrm{e}^{(a-b)x} = 0$ et la fonction $t \mapsto F(t)\mathrm{e}^{(a-b)t}$ est intégrable sur $\mathbb{R}_+$, d'après les théorèmes de comparaison et puisque $a - b < 0$. Donc $\lim\limits_{x\to+\infty} \int_0^x F(t)\mathrm{e}^{(a-b)t}\,\mathrm{d}t$ existe. On en déduit que $\lim\limits_{x\to+\infty} \int_0^x f(t)\mathrm{e}^{-bt}\,\mathrm{d}t$ existe, ce qu'il fallait démontrer.

✓ Vous noterez une fois encore l'efficacité de l'intégration par parties : elle a permis d'exprimer une intégrale semi-convergente à l'aide d'une intégrale absolument convergente. Voyez aussi l'exemple de **[SF14.4]** page 541.

SF14.10 Utiliser des inégalités, en particulier celle de Cauchy-Schwarz

L'utilisation d'inégalités pour étudier l'intégrabilité d'une fonction est évidemment fondamentale : elle est à la base de tous les théorèmes de comparaison.
Parmi les inégalités utiles, il y a évidemment celle de Cauchy-Schwarz : il faut bien sûr penser à l'utiliser dans tout exercice où intervient l'intégrale d'une fonction au carré, mais aussi celle d'un produit. Il faut même parfois faire apparaître ce produit.

Exemple

Soit $f : \mathbb{R}_+ \to \mathbb{R}$, de classe $\mathscr{C}^1$.

On suppose f et f'^2 intégrables sur $\mathbb{R}_+$. Démontrer que $\lim\limits_{x\to+\infty} f(x) = 0$.

Pour relier f à f'^2, on pense naturellement à l'inégalité de Cauchy-Schwarz. On peut en effet écrire, pour tous x, $y \in \mathbb{R}_+$:

$$|f(y) - f(x)| = \left|\int_x^y f'\right| = \left|\int_x^y 1 \times f'\right| \leqslant \left(\int_x^y 1^2\right)^{\frac{1}{2}} \left(\int_x^y f'^2\right)^{\frac{1}{2}},$$

donc : $|f(y) - f(x)| \leqslant k\sqrt{|y-x|}$, avec $k = \left(\int_{\mathbb{R}_+} f'^2\right)^{\frac{1}{2}}$.
On en déduit que f est uniformément continue sur $\mathbb{R}_+$; en effet, si on se donne $\varepsilon > 0$, en posant $\alpha = \left(\frac{\varepsilon}{k}\right)^2$, on aura bien $|x-y| \leqslant \alpha \implies |f(x) - f(y)| \leqslant \varepsilon$.
Il ne reste plus qu'à appliquer le résultat de l'exemple **[SF14.8]** *page 544 pour conclure.*

Une autre inégalité utile, que nous avons déjà employée, est la suivante :

$$\forall\,(a,b) \in \mathbb{R}^2,\ ab \leqslant \frac{1}{2}(a^2 + b^2)\,.$$

Exemple

Soit f de classe $\mathscr{C}^2$ sur $\mathbb{R}$ telle que f et f'' sont de carré intégrable.

1) Montrer que f' est de carré intégrable.

2) Montrer que $\left(\int_{\mathbb{R}} f'^2\right)^2 \leqslant \left(\int_{\mathbb{R}} f^2\right)\left(\int_{\mathbb{R}} f''^2\right)$.

1) *En intégrant par parties, on a, pour tout $x \in \mathbb{R}$:*

$$\int_0^x f'^2(t)\,\mathrm{d}t = \left[f(t)f'(t)\right]_0^x - \int_0^x f(t)f''(t)\,\mathrm{d}t \quad (*)$$

D'après l'inégalité rappelée ci-dessus :

$$\forall t \in \mathbb{R},\ |f(t)f''(t)| \leqslant \frac{1}{2}\left(f^2(t) + f''^2(t)\right),$$

d'où il résulte que ff'' est intégrable sur $\mathbb{R}$ puisque f^2 et f''^2 le sont.
La fonction $x \mapsto \displaystyle\int_0^x f'^2(t)\,\mathrm{d}t$ étant croissante, si f'^2 n'était pas intégrable sur $\mathbb{R}_+$ on aurait $\displaystyle\lim_{x\to+\infty}\int_0^x f'^2(t)\,\mathrm{d}t = +\infty$ et d'après $(*)$ on aurait aussi $\displaystyle\lim_{x\to+\infty} f(x)f'(x) = +\infty$ ce qui est impossible puisqu'alors

$$f^2(x) = f^2(0) + 2\int_0^x f(t)f'(t)\,\mathrm{d}t$$

tendrait aussi vers $+\infty$ ce qui contredit l'intégrabilité de f^2.
On a donc prouvé par l'absurde que f'^2 est intégrable sur $\mathbb{R}_+$, et on ferait de même sur $\mathbb{R}_-$ (en faisant $x \to -\infty$).

2) *De $(*)$ découle alors que $\displaystyle\lim_{x\to+\infty} f(x)f'(x) = \ell$ existe. D'autre part, ff' est intégrable sur $\mathbb{R}_+$, toujours à l'aide de l'inégalité :*

$$\forall t \in \mathbb{R}_+,\ |f(t)f'(t)| \leqslant \frac{1}{2}\left(f^2(t) + f'^2(t)\right).$$

On a donc nécessairement $\ell = 0$ (raisonner par l'absurde), et quand $x \to +\infty$ la relation $(*)$ donne :

$$\int_0^{+\infty} f'^2 = -f(0)f'(0) - \int_0^{+\infty} ff''.$$

On obtient de la même façon, lorsque $x \to -\infty$:

$$\int_0^{-\infty} f'^2 = -f(0)f'(0) - \int_0^{-\infty} ff'',$$

et en soustrayant ces deux relations : $\displaystyle\int_{\mathbb{R}} f'^2 = \int_{\mathbb{R}} ff''$.

L'inégalité de Cauchy-Schwarz donne alors le résultat souhaité.

■ Comment trouver la limite d'une intégrale dépendant d'un paramètre ?

SF14.11 Utiliser un encadrement

Dans beaucoup de cas, nul besoin d'utiliser un théorème compliqué : un simple encadrement ou une simple majoration peuvent suffire.

Exemple

Pour $x > 0$, on pose $f(x) = \displaystyle\int_0^{\frac{1}{x}} \frac{\mathrm{d}t}{x + \sin^2 t}$.

Déterminer les limites de f en 0^+ et en $+\infty$.

On remarque déjà que f est bien définie sur $\mathbb{R}_+^*$ comme intégrale d'une fonction continue sur un segment.
Puisque pour tout $x > 0$ et tout réel t, $0 < x \leqslant x + \sin^2 t \leqslant x + 1$, on a :

$$\forall\, x > 0,\ \frac{1}{x(x+1)} \leqslant f(x) \leqslant \frac{1}{x^2}\,,$$

ce qui donne $\lim\limits_{x\to 0^+} f(x) = +\infty$, $\lim\limits_{x\to+\infty} f(x) = 0$ et fournit aussi l'équivalent $f(x) \underset{x\to+\infty}{\sim} \dfrac{1}{x^2}$.

SF14.12 Utiliser le théorème de convergence dominée

Il s'agit certainement de la méthode la plus utilisée. Le théorème de convergence dominée a été rappelé en **[S14.9]**. Il faut y penser d'emblée dès que l'on a une suite d'intégrales dépendant d'un entier n, mais il peut aussi s'appliquer, à l'aide du théorème de caractérisation séquentielle de la limite, au cas des intégrales dépendant d'un paramètre réel (voir la généralisation du théorème de converegence dominée en **[S14.11]** page 149).
La seule véritable difficulté d'application de ce théorème est l'hypothèse de domination : comment trouver une fonction φ dominante pour une suite de fonctions (f_n) ? Citons quelques pistes.

- Un cas facile et assez fréquent est celui où la suite $(|f_n|)$ est monotone : dans ce cas, il conviendra de prendre $\varphi = |f_0|$ si la suite est décroissante (ou $|f_{n_0}|$ pour avoir une majoration seulement à partir d'un certain rang) , ou $\varphi = \lim\limits_{n\to+\infty} |f_n|$ si la suite est croissante. C'est de toutes façons la meilleure majoration possible.
- Pour majorer les $|f_n|$, il faut se contenter de majorer dans l'expression de $|f_n|$ la portion qui dépend de n ; on garde tous les termes qui ne dépendent pas de n, du moins dans un premier temps.
- Découper l'intervalle d'intégration en plusieurs intervalles, et faire des majorations différentes sur chacun d'entre eux ; on peut ainsi obtenir une fonction dominante seulement continue par morceaux, mais le théorème prévoit ce cas.

Exemple

Déterminer $\lim\limits_{n\to+\infty} \displaystyle\int_1^n \left(1 - \frac{x}{n}\right)^n \mathrm{d}x$.

On considère pour $n \in \mathbb{N}^*$: $f_n(x) = \begin{cases} \left(1 - \dfrac{x}{n}\right)^n & \text{si } x \in [0\,;n[\\ 0 & \text{si } x \geqslant n\,. \end{cases}$

Les fonctions f_n sont positives et continues sur $\mathbb{R}_+$. On a :

$$I_n = \int_0^n \left(1 - \frac{x}{n}\right)^n \mathrm{d}x = \int_{\mathbb{R}_+} f_n(x)\,\mathrm{d}x\,,$$

cette intégrale étant forcément convergente.
On vérifie ensuite que la suite (f_n) converge simplement sur $\mathbb{R}_+$ vers $f : x \mapsto \mathrm{e}^{-x}$ (voir exercice **12.5***).*
De plus, pour tout $x \in [0\,;n[$, $f_n(x) = \mathrm{e}^{n\ln\left(1-\frac{x}{n}\right)} \leqslant \mathrm{e}^{-x}$, en utilisant l'inégalité bien connue $\ln(1+X) \leqslant X$.

On en déduit que, pour tout $x \in \mathbb{R}_+$, $0 \leqslant f_n(x) \leqslant \mathrm{e}^{-x}$, et puisque la fonction $x \mapsto \mathrm{e}^{-x}$ est continue et intégrable sur $\mathbb{R}_+$, le théorème de convergence dominée s'applique, donc : $\displaystyle\lim_{n\to+\infty} I_n = \int_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = 1.$

✓ *Retenez la technique utilisée dans cet exemple pour pouvoir utiliser le théorème de convergence dominée lorsque l'intervalle d'intégration n'a pas des bornes fixes au départ, en prolongeant la fonction par la fonction nulle en dehors de l'intervalle initial.*

Exemple

Pour tout entier $n \geqslant 2$, on pose $I_n = \displaystyle\int_0^{+\infty} \frac{1+t^n}{\sqrt{t}+t^{2n}}\,\mathrm{d}t$.

Justifier l'existence de I_n et déterminer $\displaystyle\lim_{n\to+\infty} I_n$.

Pour $n \geqslant 2$ et $t \in \mathbb{R}_+^*$, soit $f_n(t) = \dfrac{1+t^n}{\sqrt{t}+t^{2n}}$.

Les fonctions f_n sont continues sur $\mathbb{R}_+^*$. Pour étudier la convergence de la suite $(f_n(t))$, il faut évidemment distinguer selon la position de t par rapport à 1. On obtient facilement que la suite (f_n) converge simplement sur $\mathbb{R}_+^*$ vers la fonction $f : t \mapsto \begin{cases} \dfrac{1}{\sqrt{t}} & \text{si } 0 < t < 1 \\ 1 & \text{si } t = 1 \\ 0 & \text{si } t > 1 \end{cases}$, qui est une fonction continue par morceaux sur $\mathbb{R}_+^*$.

Il reste à trouver une fonction dominante, et là aussi il est clair qu'il faut distinguer selon la position de t par rapport à 1. Plus précisément :

- *si $t \leqslant 1$, $0 \leqslant f_n(t) \leqslant \dfrac{2}{\sqrt{t}}$, en majorant t^n par 1 au numérateur et en minorant le dénominateur par $\sqrt{t}$;*
- *si $t > 1$, $0 \leqslant f_n(t) \leqslant \dfrac{2t^n}{t^{2n}} = \dfrac{2}{t^n} \leqslant \dfrac{2}{t^2}$, en majorant 1 par t^n au numérateur et en minorant le dénominateur par t^{2n}.*

On a donc finalement :

$$\forall\, n \geqslant 2,\ \forall\, t > 0,\ 0 \leqslant f_n(t) \leqslant \varphi(t)\,,$$

avec $\varphi : t \mapsto \begin{cases} \dfrac{2}{\sqrt{t}} & \text{si } 0 < t \leqslant 1 \\ \dfrac{2}{t^2} & \text{si } t > 1\,. \end{cases}$

La fonction φ est continue par morceaux (et même continue) sur $\mathbb{R}_+^*$, et elle est intégrable puisque $t \mapsto \dfrac{1}{\sqrt{t}}$ est intégrable au voisinage de 0^+ et $t \mapsto \dfrac{1}{t^2}$ l'est au voisinage de $+\infty$.

Cette majoration prouve déjà que les f_n sont bien intégrables sur $\mathbb{R}_+^*$ pour $n \geqslant 2$; de plus, le théorème de convergence dominée s'applique et la limite cherchée est égale à :

$$\int_{\mathbb{R}_+^*} f = \int_0^1 \frac{\mathrm{d}t}{\sqrt{t}} = \Big[2\sqrt{t}\Big]_0^1 = 2\,.$$

✓ *Cet exemple illustre le fait qu'il faut parfois rechercher des fonctions dominantes différentes selon l'intervalle considéré.*

Exemple

Soit $f \in \mathscr{C}([0\,;1],\mathbb{R})$. Démontrer que : $\lim\limits_{h\to 0^+} \displaystyle\int_0^1 \frac{h}{h^2+x^2} f(x)\,\mathrm{d}x = \frac{\pi}{2} f(0)$.

Il ne s'agit pas ici d'une suite de fonctions, mais de la limite d'une intégrale à paramètre. Nous utiliserons donc directement l'extension du théorème de convergence dominée rappelé en **[S14.11]**.
On voit que celui-ci ne peut s'appliquer directement (sinon, en intervertissant limite et intégrale, on obtiendrait 0 pour limite). Il nous faut donc dans un premier temps transformer l'intégrale, en faisant le changement de variable affine $x = ht$ pour $h > 0$:

$$\forall\, h > 0,\ I(h) = \int_0^1 \frac{h}{h^2+x^2} f(x)\,\mathrm{d}x = \int_0^{1/h} \frac{f(ht)}{1+t^2}\,\mathrm{d}t\,.$$

Considérons alors la fonction g définie sur $\mathbb{R}_+^* \times \mathbb{R}_+$ par :

$$g(h,t) = \begin{cases} \dfrac{f(ht)}{1+t^2} & \text{si } 0 \leqslant t \leqslant \dfrac{1}{h} \\ 0 & \text{si } t > \dfrac{1}{h} \end{cases},$$

de sorte que : $\forall\, h > 0,\ I(h) = \displaystyle\int_0^{+\infty} g(h,t)\,\mathrm{d}t$.

- $\forall\, h > 0$, *la fonction* $t \mapsto g(h,t)$ *est continue par morceaux sur* $\mathbb{R}_+$.
- *Pour tout* $t \geqslant 0$, $\lim\limits_{h\to 0^+} g(h,t) = \ell(t) = \dfrac{f(0)}{1+t^2}$ *; en effet, tout réel t positif est compris entre 0 et $\dfrac{1}{h}$ pour h assez petit, et* $\lim\limits_{h\to 0^+} f(ht) = f(0)$ *puisque f est continue. De plus, la fonction limite ℓ est continue (par morceaux).*
- *f étant continue sur $[0\,;1]$ y est bornée et :*

$$\forall\, h > 0,\ \forall\, t \in \mathbb{R}_+,\ |g(h,t)| \leqslant \frac{\|f\|_\infty}{1+t^2} = \varphi(t)\,,$$

où φ est continue et intégrable sur $\mathbb{R}_+$.

Le théorème de convergence dominée généralisé s'applique donc et :

$$\lim_{h\to 0^+} I(h) = \int_0^{+\infty} \frac{f(0)}{1+t^2}\,\mathrm{d}t = \frac{\pi}{2} f(0)\,.$$

✓ *Notez encore une fois la méthode permettant de ramener l'intégrale sur un intervalle dont les bornes dépendent de h à une intégrale sur un intervalle fixe pour pouvoir appliquer le théorème de convergence dominée.*

SF14.13 Conjecturer la limite

Il arrive souvent, même lorsque le théorème de convergence dominée ne s'applique pas, que la limite de l'intégrale soit égale à l'intégrale de la limite (pour simplifier). On peut donc essayer de « deviner » la limite, puis former la différence et montrer

rigoureusement que cette différence tend vers 0. Lorsque la limite conjecturée est infinie, on peut aussi utiliser un raisonnement par l'absurde.

Exemple

Soit f continue de $\mathbb{R}_+$ dans $\mathbb{R}$ et, pour $n \geqslant 1$,$u_n = n\displaystyle\int_1^{1+\frac{1}{n}} f(x^n)\,\mathrm{d}x$.
Déterminer $\lim\limits_{n\to+\infty} u_n$.

Un examen rapide de l'énoncé doit déjà vous convaincre qu'aucune méthode classique ne peut ici s'appliquer directement.
Le changement de variable $t = x^n$ pour $n \geqslant 1$ s'impose ! Il conduit (puisque $x = t^{\frac{1}{n}}$, $\mathrm{d}x = \frac{1}{n}t^{\frac{1}{n}-1}\,\mathrm{d}t$) à :

$$u_n = \int_1^{x_n} t^{\frac{1}{n}}\frac{f(t)}{t}\,\mathrm{d}t \quad \text{avec } x_n = \left(1+\frac{1}{n}\right)^n.$$

Puisque

$$\lim_{n\to+\infty} t^{\frac{1}{n}} = \lim_{n\to+\infty} \mathrm{e}^{\frac{1}{n}\ln t} = 1 \quad \text{pour } t > 0\,,$$

et que

$$\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n = \lim_{n\to+\infty}\mathrm{e}^{n\ln\left(1+\frac{1}{n}\right)} = \mathrm{e}\,,$$

on s'attend naturellement à ce que la limite de l'intégrale ci-dessus soit égale à $\ell = \displaystyle\int_1^{\mathrm{e}}\frac{f(t)}{t}\,\mathrm{d}t$. Pour le démontrer, on forme la différence :

$$u_n - \ell = \int_1^{x_n} t^{\frac{1}{n}}\frac{f(t)}{t}\,\mathrm{d}t - \int_1^{\mathrm{e}}\frac{f(t)}{t}\,\mathrm{d}t = \int_1^{x_n}\left(t^{\frac{1}{n}}-1\right)\frac{f(t)}{t}\,\mathrm{d}t - \int_{x_n}^{\mathrm{e}}\frac{f(t)}{t}\,\mathrm{d}t\,.$$

– *Puisque que la fonction $t \mapsto \dfrac{f(t)}{t}$ est continue sur $[1\,;\mathrm{e}]$,*

$$\lim_{n\to+\infty}\int_{x_n}^{\mathrm{e}}\frac{f(t)}{t}\,\mathrm{d}t = \int_{\mathrm{e}}^{\mathrm{e}}\frac{f(t)}{t}\,\mathrm{d}t = 0\,.$$

– *D'autre part, pour tout $x \geqslant 0$ on a, d'après le théorème des accroissements finis appliqué à la fonction* exp *sur $[0\,;x]$:*

$$0 \leqslant \mathrm{e}^x - 1 \leqslant x\mathrm{e}^x$$

donc, avec $x = \frac{1}{n}\ln t$: $\forall\, t \in [1\,;\mathrm{e}]$, $0 \leqslant t^{\frac{1}{n}} - 1 \leqslant \frac{\ln t}{n}t^{\frac{1}{n}} \leqslant t\frac{\ln t}{n} \leqslant \frac{t}{n}$,
d'où :

$$\left|\int_1^{x_n}\left(t^{\frac{1}{n}}-1\right)\frac{f(t)}{t}\,\mathrm{d}t\right| \leqslant \frac{1}{n}\int_1^{x_n}|f(t)|\,\mathrm{d}t \leqslant \frac{1}{n}\int_1^{\mathrm{e}}|f(t)|\,\mathrm{d}t$$

(car $x_n \leqslant \mathrm{e}$ pour tout n). Il en résulte que :

$$\lim_{n\to+\infty}\left(\int_1^{x_n}\left(t^{\frac{1}{n}}-1\right)\frac{f(t)}{t}\,\mathrm{d}t\right) = 0\,,$$

puis : $\lim\limits_{n\to+\infty} u_n = \ell$.

Exemple

Pour $x \in\,]-1\,;1[$ on pose : $f(x) = \displaystyle\int_0^{\frac{\pi}{2}}\frac{\mathrm{d}t}{\sqrt{1-x^2\sin^2 t}}$.

Déterminer la limite de $f(x)$ quand $x \mapsto 1^-$.

On remarque déjà que pour $x \in]-1\,;1[$ on a $1 - x^2\sin^2 t > 0$ pour tout t, donc l'intégrale existe.

On voudrait pouvoir dire que la limite de l'intégrale est l'intégrale de la limite, c'est-à-dire ici $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{\cos t}$. Mais cette intégrale est divergente (cela sera démontré un peu plus loin) !

Cela suggère donc de montrer que $\lim\limits_{x\to 1^-} f(x) = +\infty$.

Pour cela, on remarque déjà que f est croissante sur $[0\,;1[$; en effet, si $0 \leqslant x \leqslant y < 1$, on aura $x^2 \leqslant y^2 < 1$ puis $1 - x^2\sin^2 t \geqslant 1 - y^2\sin^2 t$ pour tout t donc $f(x) \leqslant f(y)$.

D'après le théorème de la limite monotone, dire que $\lim\limits_{x\to 1^-} f(x) = +\infty$ équivaut à dire que f n'est pas majorée.

Raisonnons donc par l'absurde en supposant qu'il existe un réel M tel que $\forall\, x \in [0\,;1[,\ f(x) \leqslant M$.

On aurait alors :

$$\forall\, x \in [0\,;1[,\ \forall\, a \in \left[0\,;\tfrac{\pi}{2}\right[,\ \int_0^a \frac{\mathrm{d}t}{\sqrt{1 - x^2\sin^2 t}} \leqslant M \quad (*)$$

Mais lorsque $t \in [0\,;a]$ avec $a < \frac{\pi}{2}$, les fonctions $t \mapsto \dfrac{1}{\sqrt{1 - x^2\sin^2 t}}$ sont dominées indépendamment de x par la fonction $t \mapsto \dfrac{1}{\sqrt{1 - \sin^2 t}} = \dfrac{1}{\cos t}$ qui est intégrable sur $[0\,;a]$. Le théorème de continuité d'une intégrale à paramètre permet alors d'écrire :

$$\lim_{x\to 1^-} \int_0^a \frac{\mathrm{d}t}{\sqrt{1 - x^2\sin^2 t}} = \int_0^a \frac{\mathrm{d}t}{\cos t}\,.$$

En passant à la limite dans $(*)$ cela implique :

$$\forall\, a \in \left[0\,;\tfrac{\pi}{2}\right[,\ \int_0^a \frac{\mathrm{d}t}{\cos t} \leqslant M\,.$$

Mais la fonction $t \mapsto \dfrac{1}{\cos t}$ étant positive, cette majoration signifie que l'intégrale $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{\cos t}$ serait convergente ; or cela n'est pas le cas puisque :

$$\int_0^x \frac{\mathrm{d}t}{\cos t} = \ln\left|\tan\left(\frac{x}{2} + \frac{\pi}{4}\right)\right| \underset{x\to\frac{\pi}{2}^-}{\longrightarrow} +\infty\,.$$

Il y a donc contradiction : f ne peut être majorée sur $[0\,;1[$ donc, étant croissante, $\lim\limits_{x\to 1^-} f(x) = +\infty$.

SF14.14 Utiliser une approximation

L'idée est ici de remplacer la fonction à intégrer par une fonction équivalente plus simple puis de montrer que la différence des deux expressions tend vers 0.

Exemple

Calculer $\displaystyle\lim_{x\to 1} \int_x^{x^2} \frac{\mathrm{d}t}{\ln t}$.

- *Il faut bien sûr déjà vérifier que l'expression de l'énoncé a bien un sens. Pour $x \in]0\,;1[$ on a $0 < x^2 < x < 1$, donc la fonction $t \mapsto \dfrac{1}{\ln t}$ est bien continue sur l'intervalle $\left[x^2\,;x\right]$, et l'intégrale existe.*
  On démontre de même son existence lorsque $x > 1$. Ainsi, la fonction $x \mapsto \displaystyle\int_x^{x^2} \frac{\mathrm{d}t}{\ln t}$ est-elle définie sur $\mathbb{R}_+^* \setminus \{1\}$.
- *On connaît l'équivalent $\ln t \underset{t\to1}{\sim} t-1$, donc $\dfrac{1}{\ln t} \underset{t\to1}{\sim} \dfrac{1}{t-1}$.*
  Considérons alors la fonction $\varphi\colon t \mapsto \dfrac{1}{\ln t} - \dfrac{1}{t-1}$. Cette fonction est continue sur $\mathbb{R}_+^* \setminus \{1\}$; de plus, en posant $t = 1+h$ et en utilisant les développements limités usuels on a :
  $$\varphi(1+h) = \frac{1}{\ln(1+h)} - \frac{1}{h} = \frac{h-\ln(1+h)}{h\ln(1+h)} \underset{h\to0}{\sim} \frac{h^2/2}{h^2}$$
  donc φ se prolonge en une fonction continue sur $\mathbb{R}_+^*$ en posant $\varphi(1) = \dfrac{1}{2}$.
- *On a alors pour tout $x \in \mathbb{R}_+^* \setminus \{1\}$:*
  $$\int_x^{x^2} \frac{\mathrm{d}t}{\ln t} = \int_x^{x^2} \frac{\mathrm{d}t}{t-1} + \int_x^{x^2} \varphi(t)\,\mathrm{d}t = \ln\left(\frac{x^2-1}{x-1}\right) + \int_x^{x^2} \varphi(t)\,\mathrm{d}t\,.$$
  φ étant continue sur $\mathbb{R}_+^*$, $\displaystyle\lim_{x\to1}\int_x^{x^2} \varphi(t)\,\mathrm{d}t = \int_1^1 \varphi(t)\,\mathrm{d}t = 0$ donc :
  $$\lim_{x\to1}\int_x^{x^2} \frac{\mathrm{d}t}{\ln t} = \lim_{x\to1} \ln(x+1) = \ln 2\,.$$

  ✓ *On a utilisé ici une propriété vue en cours, et qui est importante pour certains passages à la limite : l'intégrale d'une fonction continue est une fonction continue de ses bornes.*

SF14.15 Utiliser le théorème de continuité d'une intégrale à paramètre

Si l'expression qui est donnée est une fonction continue d'un paramètre x, sa limite en un point n'est rien d'autre que sa valeur en ce point.

Exemple

Déterminer $\displaystyle\lim_{x\to0^+}\int_0^1 \frac{1-t^x}{1-t}\,\mathrm{d}t$.

Pour $(x,t) \in \mathbb{R}_+ \times]0\,;1[$ posons $f(x,t) = \dfrac{1-t^x}{1-t} = \dfrac{1-\mathrm{e}^{x\ln t}}{1-t}$.
f est continue par rapport à x et par rapport à t ; de plus, pour $x \in [0\,;1]$ et $t \in]0\,;1[$ on a $1 \geqslant t^x \geqslant t$ d'où $0 \leqslant f(x,t) \leqslant 1$.
L'application constante égale à 1 étant continue et intégrable sur $]0\,;1[$, f vérifie l'hypothèse de domination sur $[0\,;1] \times]0\,;1[$.
Il résulte du théorème de continuité d'une intégrale dépendant d'un paramètre que la fonction $x \mapsto \displaystyle\int_0^1 f(x,t)\,\mathrm{d}t$ est continue sur $[0\,;1]$. En particulier elle est continue en 0^+ donc :
$$\lim_{x\to0^+}\int_0^1 \frac{1-t^x}{1-t}\,\mathrm{d}t = \int_0^1 f(0,t)\,\mathrm{d}t = 0\,.$$

SF14.16 Se ramener à un calcul de limite

Par un changement de variable ou une intégration par parties, on peut souvent faire apparaître un terme intéressant puis déterminer un équivalent en cherchant alors la limite d'une intégrale, ce qui nous ramène au savoir-faire précédent.

Exemple

On pose $f(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t$.

Déterminer le domaine de définition de f et trouver un équivalent de $f(x)$ quand x tend vers $+\infty$.

- – *Si $x < 0$, la fonction $t \mapsto \dfrac{\mathrm{e}^{-t}}{x+t}$ n'est pas définie sur $]0\,;+\infty[$ et $f(x)$ n'existe pas.*
  – *Si $x \geqslant 0$, $\dfrac{\mathrm{e}^{-t}}{x+t} \underset{t\to 0}{\sim} \dfrac{1}{x+t}$ donc est intégrable au voisinage de 0 si et seulement si $x > 0$.*
  – *De plus, $\dfrac{\mathrm{e}^{-t}}{x+t} = o(\mathrm{e}^{-t})$ quand $t \to +\infty$, et $t \mapsto \mathrm{e}^{-t}$ est intégrable sur $\mathbb{R}_+$, donc $t \mapsto \dfrac{\mathrm{e}^{-t}}{x+t}$ est intégrable au voisinage de $+\infty$.*
  En conclusion, le domaine de définition de f est $\mathbb{R}_+^*$.
- *On écrit $f(x) = \dfrac{1}{x}\displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-t}}{1+\frac{t}{x}}\,\mathrm{d}t$ puis $\displaystyle\lim_{x\to+\infty} \int_0^{+\infty} \frac{\mathrm{e}^{-t}}{1+\frac{t}{x}}\,\mathrm{d}t = \int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t$ par convergence dominée (nous vous laissons le soin de détailler), donc $f(x) \underset{x\to+\infty}{\sim} \dfrac{1}{x}$.*

Exemple

Pour tout $n \in \mathbb{N}^*$, on pose $u_n = \displaystyle\int_0^{+\infty} \frac{\cos t}{1+n^2t^2}\,\mathrm{d}t$.

Justifier l'existence de u_n, et en déterminer la limite puis un équivalent quand $n \to +\infty$.

- *Pour tout $n \in \mathbb{N}^*$, la fonction $f_n : t \mapsto \dfrac{\cos t}{1+n^2t^2}$ est continue sur $[0\,;+\infty[$, et on a pour tout $t \in \mathbb{R}_+$, $|f_n(t)| \leqslant \dfrac{1}{1+t^2}$ qui est continue positive et intégrable sur $\mathbb{R}_+$. Cela prouve l'intégrabilité de f_n.*
- *La suite $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}_+$ vers $f : t \mapsto \begin{cases} 1 & \text{si } t = 0 \\ 0 & \text{si } t > 0 \end{cases}$. L'hypothèse de domination a déjà été démontrée plus haut donc le théorème de convergence dominée s'applique : $\displaystyle\lim_{n\to+\infty} u_n = \int_0^{+\infty} f = 0$.*
- *Pour obtenir un équivalent, le changement de variable $u = nt$ s'impose. Il permet d'écrire : $u_n = \dfrac{1}{n}\displaystyle\int_0^{+\infty} \frac{\cos\left(\frac{u}{n}\right)}{1+u^2}\,\mathrm{d}u$.*

Là encore le théorème de convergence dominée permet de montrer que :

$$\lim_{n\to+\infty}\int_0^{+\infty}\frac{\cos\left(\frac{u}{n}\right)}{1+u^2}\,\mathrm{d}u=\int_0^{+\infty}\frac{\mathrm{d}u}{1+u^2}=\frac{\pi}{2}\,,$$

ce qui permet de conclure : $u_n \underset{n\to+\infty}{\sim} \dfrac{\pi}{2n}$.

Exemple

On pose, pour $n \geqslant 1$, $I_n = \displaystyle\int_0^1 \frac{1-x^{1/n}}{1-x}\,\mathrm{d}x$. Déterminer la limite de I_n et un équivalent quand $n \to +\infty$.

- *Soit, pour* $n \geqslant 1$ *et* $x \in [0\,;1[$, $f_n(x) = \dfrac{1-x^{1/n}}{1-x}$. f_n *est continue sur* $[0\,;1[$, *et* $\lim\limits_{x\to1^-} f_n(x) = \dfrac{1}{n}$ *(cette limite est en effet la dérivée de la fonction* $x \mapsto x^{1/n}$ *en* 1*). Ainsi,* f_n *se prolonge par continuité sur* $[0\,;1]$, *et y est par suite intégrable.*
- *La suite* f_n *(prolongée en* 1*) converge simplement sur* $[0\,;1]$ *vers la fonction* $f : x \mapsto \begin{cases} 0 & \text{si } x \in]0\,;1] \\ 1 & \text{si } x = 0 \end{cases}$.

  Remarquons que f *n'étant pas continue alors que les* f_n *le sont, la convergence n'est pas uniforme ; on ne peut donc pas appliquer directement le théorème de passage à la limite vu en* **[S12.2]**.

  Pour tout $x \in [0\,;1]$, *on a* $0 \leqslant f_n(x) \leqslant 1$. *La fonction constante* $x \mapsto 1$ *étant continue et intégrable sur* $[0\,;1]$, *le théorème de convergence dominée s'applique, ce qui donne* $\lim\limits_{n\to+\infty} I_n = \displaystyle\int_0^1 f = 0$.
- *Effectuons une intégration par parties sur un segment* $[\varepsilon,a]$ *avec* $0 < \varepsilon < a < 1$ *:*

  $$\int_\varepsilon^a \frac{1-x^{1/n}}{1-x}\,\mathrm{d}x = \Big[-(1-x^{1/n})\ln(1-x)\Big]_\varepsilon^a - \frac{1}{n}\int_\varepsilon^a x^{1/n-1}\ln(1-x)\,\mathrm{d}x\,.$$

  Or $\lim\limits_{a\to1^-}(1-a^{1/n})\ln(1-a) = 0$ *(car* $1-x^{1/n} \underset{x\to1}{\sim} \frac{1}{n}(1-x)$*) et* $\lim\limits_{\varepsilon\to0}(1-\varepsilon^{1/n})\ln(1-\varepsilon) = 0$, *donc*

  $$I_n = -\frac{1}{n}\int_0^1 x^{1/n-1}\ln(1-x)\,\mathrm{d}x\,.$$

  L'utilisation du théorème de convergence dominée permet là encore de montrer que :

  $$\lim_{n\to+\infty}\int_0^1 x^{1/n-1}\ln(1-x)\,\mathrm{d}x = \int_0^1 \frac{\ln(1-x)}{x}\,\mathrm{d}x$$

  (nous ne détaillons pas, mais nous vous invitons à en vérifier soigneusement les hypothèses).

  En notant $J = -\displaystyle\int_0^1 \frac{\ln(1-x)}{x}\,\mathrm{d}x$, *on a donc :* $I_n \underset{n\to\infty}{\sim} \frac{1}{n}J$. *Reste à calculer* J, *ce qui est assez classique.*

Un développement en série entière donne

$$\forall\, x \in [0\,;1[,\ -\frac{\ln(1-x)}{x} = \sum_{n=1}^{+\infty} \frac{x^{n-1}}{n}\cdot$$

Le théorème d'intégration terme à terme (en vérifier soigneusement les hypothèses...) donne alors : $J = \displaystyle\sum_{n=1}^{+\infty} \int_0^1 \frac{x^{n-1}}{n}\,\mathrm{d}x = \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}\cdot$

Finalement, on a : $I_n \underset{n\to+\infty}{\sim} \dfrac{\pi^2}{6n}\cdot$

SF14.17 Former la différence

Si l'équivalent à trouver est donné par l'énoncé, ou s'il est facile à deviner, on peut former la différence entre les deux quantités et montrer que la quantité obtenue est négligeable par rapport à l'un des deux termes.

Exemple

Soient f et g deux fonctions continues sur $\mathbb{R}_+$, à valeurs positives. On suppose f intégrable et g bornée, et l'on note, pour tout $x > 0$, $\varphi(x) = \displaystyle\int_0^{+\infty} \frac{f(t)}{x+g(t)}\,\mathrm{d}t$.

Montrer que φ est bien définie sur $\mathbb{R}_+^*$ et que $\varphi(x) \underset{x\to+\infty}{\sim} \dfrac{1}{x}\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t$.

- *L'existence de φ ne pose pas de problème, puisque, pour $x > 0$ fixé, l'on a* $\left|\dfrac{f(t)}{x+g(t)}\right| \leqslant \dfrac{1}{x}\,|f(t)|$ *avec f intégrable.*
- *On forme la différence :*

$$\begin{aligned}\left|\varphi(x) - \frac{1}{x}\int_0^{+\infty} f(t)\,\mathrm{d}t\right| &= \left|\int_0^{+\infty}\left(\frac{f(t)}{x+g(t)} - \frac{f(t)}{x}\right)\mathrm{d}t\right| \\ &= \left|\int_0^{+\infty} \frac{-f(t)g(t)}{x(x+g(t))}\,\mathrm{d}t\right| \\ &\leqslant \frac{\|g\|_\infty}{x}\int_0^{+\infty}\frac{f(t)}{x+g(t)}\,\mathrm{d}t\,,\end{aligned}$$

puisque toutes les fonctions sont positives, et ainsi :

$$\varphi(x) - \frac{1}{x}\int_0^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} o(\varphi(x))\ ,$$

ce qui prouve le résultat demandé.

✓ Un rappel important : dire que deux fonctions u et v sont équivalentes au voisinage d'un point signifie : $u - v = o(u)$ (ou $o(v)$). Dans les exercices un peu théoriques, cette caractérisation est souvent plus pratique à utiliser et plus utile que $\lim \dfrac{u}{v} = 1$.

SF14.18 Utiliser une intégration par parties

Cette méthode ressemble à la précédente : en faisant une intégration par parties, on peut écrire l'intégrale comme somme de deux quantités ; si l'une de ces quantités est négligeable par rapport à l'autre, on obtiendra un équivalent.

Exemple

Déterminer un équivalent de $\displaystyle\int_x^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$ lorsque x tend vers $+\infty$.

La convergence de l'intégrale proposée est facile à établir.
On réalise ensuite une intégration par parties. Pour $x>0$ on a :

$$\int_x^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \int_x^{+\infty} \left(2t\mathrm{e}^{-t^2}\right)\left(\frac{1}{2t}\right)\mathrm{d}t$$
$$= \left[-\frac{\mathrm{e}^{-t^2}}{2t}\right]_x^{+\infty} - \frac{1}{2}\int_x^{+\infty} \frac{\mathrm{e}^{-t^2}}{t^2}\,\mathrm{d}t = \frac{\mathrm{e}^{-x^2}}{2x} - \frac{1}{2}\int_x^{+\infty} \frac{\mathrm{e}^{-t^2}}{t^2}\,\mathrm{d}t\,,$$

cette intégration par parties étant justifiée puisque le terme entre crochets a bien une limite (nulle) en $+\infty$.
Or $0 \leqslant \displaystyle\int_x^{+\infty} \frac{\mathrm{e}^{-t^2}}{t^2}\,\mathrm{d}t \leqslant \frac{1}{x^2}\int_x^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$ *donc quand* $x\to+\infty$,

$$\int_x^{+\infty} \frac{\mathrm{e}^{-t^2}}{t^2}\,\mathrm{d}t = o\left(\int_x^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t\right),$$

ce qui prouve que :

$$\int_x^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^{-x^2}}{2x}\cdot$$

SF14.19 Découper l'intervalle

Cette technique est délicate et demande pas mal d'intuition. Elle consiste à découper l'intervalle d'intégration en deux parties afin d'isoler le point qui pose problème lors du passage à la limite.

Exemple

On pose $f(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t$ (on a vu dans l'exemple page 554 que f est bien définie sur $\mathbb{R}_+^*$). Trouver un équivalent de $f(x)$ lorsque $x\to 0^+$.

On effectue un développement limité de $t\mapsto \mathrm{e}^{-t}$ au voisinage de 0 : $\mathrm{e}^{-t} = 1-t+t\varphi(t)$ *avec* $\lim\limits_{t\to 0}\varphi(t)=0$. *$\varphi$ est donc prolongeable par continuité sur un voisinage de 0 de la forme $[0\,;A]$ avec $A>0$. On a alors :*

$$f(x) = \int_0^A \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t + \int_A^{+\infty} \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t$$
$$= \int_0^A \frac{1-t}{x+t}\,\mathrm{d}t + \int_0^A \frac{t\varphi(t)}{x+t}\,\mathrm{d}t + \int_A^{+\infty} \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t\,.$$

Or pour tout $x>0$:

$$\left|\int_0^A \frac{t\varphi(t)}{x+t}\,\mathrm{d}t + \int_A^{+\infty} \frac{\mathrm{e}^{-t}}{x+t}\,\mathrm{d}t\right| \leqslant \int_0^A |\varphi(t)|\,\mathrm{d}t + \int_A^{+\infty} \frac{\mathrm{e}^{-t}}{A}\,\mathrm{d}t$$
$$\leqslant \int_0^A |\varphi(t)|\,\mathrm{d}t + \frac{\mathrm{e}^{-A}}{A}$$

et aussi : $\displaystyle\int_0^A \frac{1-t}{x+t}\,\mathrm{d}t = \int_0^A \frac{(1+x)-(x+t)}{x+t}\,\mathrm{d}t = (1+x)\ln\left(\frac{x+A}{x}\right) - A\,.$

Finalement, $f(x) = (1+x)\ln\left(\dfrac{x+A}{x}\right) + O(1) = -\ln x + O(1)$ *d'où l'on tire :* $f(x) \underset{x\to 0}{\sim} -\ln x$.

SF14.20 Utiliser le théorème de dérivation d'une intégrale à paramètre

Exemple

Pour x réel, on pose : $f(x) = \displaystyle\int_0^{\frac{\pi}{2}} \frac{\sin xt}{\sin t}\,\mathrm{d}t$.

Justifier l'existence de $f(x)$ et et déterminer un équivalent de $f(x)$ lorsque $x \to 0$.

- *Nous allons directement démontrer que* f *est de classe* $\mathscr{C}^1$ *sur* $\mathbb{R}$ *en utilisant le théorème sur la dérivabilité d'une intégrale à paramètre.*

  Pour cela, notons $g(x,t) = \dfrac{\sin xt}{\sin t}$ *pour* $(x,t) \in \mathbb{R}_+ \times \left]0\,;\frac{\pi}{2}\right]$. *On vérifie alors les hypothèses du théorème.*

  – *Soit* $x \in \mathbb{R}$. *La fonction* $t \mapsto g(x,t)$ *est continue sur* $\left]0\,;\frac{\pi}{2}\right]$ *; de plus,*

  $$\lim_{t\to 0} g(x,t) = \lim_{t\to 0} \frac{\sin xt}{\sin t} = \lim_{t\to 0} \frac{xt}{t} = x$$

  donc la fonction $t \mapsto g(x,t)$ *se prolonge par continuité en* 0 *et est par suite intégrable sur* $\left]0\,;\frac{\pi}{2}\right]$.

  – $\dfrac{\partial g}{\partial x} : (x,t) \mapsto \dfrac{t\cos xt}{\sin t}$ *existe et est continue par rapport à* x *et à* t *sur* $\mathbb{R} \times \left]0\,;\frac{\pi}{2}\right]$.

  – $\forall\,(x,t) \in \mathbb{R} \times \left]0\,;\frac{\pi}{2}\right]$, *on a :* $\left|\dfrac{\partial g}{\partial x}(x,t)\right| = \left|\dfrac{t\cos xt}{\sin t}\right| \leqslant \dfrac{t}{\sin t}$, *et la fonction* $t \mapsto \dfrac{t}{\sin t}$ *est continue sur* $\left]0\,;\frac{\pi}{2}\right]$, *et intégrable car prolongeable par continuité en* 0.

  L'hypothèse de domination est donc vérifiée.

  On en conclut que f *est de classe* $\mathscr{C}^1$ *sur* $\mathbb{R}$ *et que*

  $$\forall\, x \in \mathbb{R},\ f'(x) = \int_0^{\frac{\pi}{2}} \frac{t\cos xt}{\sin t}\,\mathrm{d}t\,.$$

- *On a alors le développement limité en* 0 *:*

  $$f(x) = f(0) + xf'(0) + o(x)$$

  et puisque $f(0) = 0$, *on a :* $f(x) \underset{x\to 0}{\sim} xf'(0)$, *avec* $f'(0) = \displaystyle\int_0^{\frac{\pi}{2}} \frac{t}{\sin t}\,\mathrm{d}t \neq 0$ *(il faut préciser* $f'(0) \neq 0$, *sinon l'équivalent précédent n'a pas de sens).*

SF14.21 Utiliser le cours : intégration des relations de comparaison

Les théorèmes rappelés en **[S14.8]** permettent parfois d'obtenir directement un équivalent, quitte à transformer un peu l'intégrale proposée auparavant.

Exemple

Montrer que : $\displaystyle\int_0^1 \frac{\mathrm{e}^{-xt}}{\sqrt{1-t}}\,\mathrm{d}t \underset{x\to+\infty}{\sim} \frac{1}{x}$.

- *Montrons déjà que l'intégrale proposée existe.*
  Pour tout $x \in \mathbb{R}$, la fonction $t \mapsto \dfrac{\mathrm{e}^{-xt}}{\sqrt{1-t}}$ est continue sur $[0\,;1[$, et est équivalente quand $x \to 1^-$ à $\dfrac{\mathrm{e}^{-x}}{\sqrt{1-t}}$ qui est une fonction de signe constant et intégrable au voisinage de 1^- (fonction de Riemann avec $\alpha = \dfrac{1}{2}$).
- *Notons $f(x)$ l'intégrale proposée. Par le changement de variable (classique) $u = \sqrt{1-t}$, qui réalise une bijection de classe $\mathscr{C}^1$ de $[0\,;1[$ sur $]0\,;1]$, on obtient :*

$$f(x) = \int_0^1 \frac{\mathrm{e}^{-xt}}{\sqrt{1-t}}\,\mathrm{d}t = -\int_1^0 \frac{\mathrm{e}^{-x(1-u^2)}}{u}(-2u)\,\mathrm{d}u = 2\mathrm{e}^{-x}\int_0^1 \mathrm{e}^{xu^2}\,\mathrm{d}u\,,$$

puis, en se limitant au cas $x \geqslant 0$, par le changement de variable affine $v = \sqrt{x}u$ on obtient $f(x) = \dfrac{2\mathrm{e}^{-x}}{\sqrt{x}}\displaystyle\int_0^{\sqrt{x}} \mathrm{e}^{v^2}\,\mathrm{d}v$.

La fonction $v \mapsto \mathrm{e}^{v^2}$ n'est pas intégrable sur $\mathbb{R}_+$. On pourra donc appliquer le théorème d'intégration des relations de comparaison à condition de trouver un équivalent intéressant de e^{v^2} lorsque $v \to +\infty$.
Or si l'on pose $g(v) = \dfrac{\mathrm{e}^{v^2}}{v}$, on a pour $v > 0$, $g'(v) = 2\mathrm{e}^{v^2} - \dfrac{\mathrm{e}^{v^2}}{v^2}$ donc $g'(v) \underset{v\to+\infty}{\sim} 2\mathrm{e}^{v^2}$ et par conséquent :

$$\int_0^{\sqrt{x}} \mathrm{e}^{v^2}\,\mathrm{d}v \underset{x\to+\infty}{\sim} \int_1^{\sqrt{x}} \mathrm{e}^{v^2}\,\mathrm{d}v \underset{x\to+\infty}{\sim} \int_1^{\sqrt{x}} \frac{g'(v)}{2}\,\mathrm{d}v = \frac{1}{2}\left(g(\sqrt{x}) - g(1)\right) \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^x}{2\sqrt{x}}\,,$$

et en rassemblant ces résultats, on obtient l'équivalent de $f(x)$ demandé.

■ Comment étudier une intégrale dépendant d'un paramètre ?

Contrairement aux deux problématiques précédentes (recherche de limite ou d'équivalent) qui peuvent donner lieu à des problèmes délicats, l'étude de la continuité et/ou de la dérivabilité d'une intégrale dépendant d'un paramètre, et éventuellement son calcul, ne nécessite le plus souvent que la bonne connaissance et l'application rigoureuse des théorèmes du cours rappelés en **[S14.11]**.
Seule l'hypothèse de domination peut être délicate à vérifier ; si il y a des difficultés, il ne faut pas oublier que l'on peut se restreindre à des segments, la continuité et la dérivabilité étant des notions locales.

SF14.22 Utiliser les théorèmes du cours

Nous donnons ci-dessous quelques exemples d'application de ces théorèmes ; ces exemples détaillés sont à étudier soigneusement car ils sont très représentatifs des problèmes qui peuvent être posés.

Exemple

Calculer $g\colon x \mapsto \displaystyle\int_0^{\frac{\pi}{2}} \ln(x\cos^2 t + \sin^2 t)\,\mathrm{d}t$ pour $x > 0$.

En déduire la valeur de l'intégrale $\displaystyle\int_0^{\frac{\pi}{2}} \ln(\sin t)\,\mathrm{d}t$.

L'idée est ici de calculer g' puis d'en déduire g par intégration, puis de faire un passage à la limite. Tout cela doit bien sûr être soigneusement justifié.

On notera $f(x,t) = \ln(x\cos^2 t + \sin^2 t)$.

- Dérivabilité sur $\mathbb{R}_+^*$

  Nous appliquons ici directement le théorème sur la dérivabilité d'une intégrale à paramètre.

  – *Pour tout $x > 0$ fixé, l'application $t \mapsto f(x,t)$ est continue sur le segment $\left[0\,;\frac{\pi}{2}\right]$, puisque $x\cos^2 t + \sin^2 t$ est alors strictement positif. Elle y est donc intégrable.*

  – $\dfrac{\partial f}{\partial x}(x,t) = \dfrac{\cos^2 t}{x\cos^2 t + \sin^2 t}$ *donc* $\dfrac{\partial f}{\partial x}(x,t)$ *est définie et continue sur* $\mathbb{R}_+^* \times \left[0\,;\frac{\pi}{2}\right]$.

  – *Soit a un réel > 0. Pour tout $x \in [a\,;+\infty[$ et tout $t \in \left[0\,;\frac{\pi}{2}\right]$ on a :*

  $$\left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \frac{\cos^2 t}{a\cos^2 t + \sin^2 t} = \varphi(t)$$

  avec φ continue donc intégrable sur le segment $\left[0\,;\frac{\pi}{2}\right]$.

  D'après le théorème du cours, g est de classe $\mathscr{C}^1$ sur tout intervalle $[a\,;+\infty[$ donc sur $\mathbb{R}_+^*$, et : $\forall\, x > 0,\ g'(x) = \displaystyle\int_0^{\frac{\pi}{2}} \frac{\cos^2 t}{x\cos^2 t + \sin^2 t}\,\mathrm{d}t$.

  En posant $u = \tan t$ (bijection strictement monotone de $\left[0\,;\frac{\pi}{2}\right[$ sur $[0\,;+\infty[$, de classe $\mathscr{C}^1$), on obtient, pour $x \neq 1$:

  $$g'(x) = \int_0^{+\infty} \frac{\mathrm{d}u}{(1+u^2)(x+u^2)} = \frac{1}{x-1}\int_0^{+\infty}\left(\frac{1}{1+u^2} - \frac{1}{x+u^2}\right)\mathrm{d}u$$
  $$= \frac{1}{x-1}\left[\operatorname{Arctan} u - \frac{1}{\sqrt{x}}\operatorname{Arctan}\frac{u}{\sqrt{x}}\right]_0^{+\infty} = \frac{\pi}{2}\frac{1}{x+\sqrt{x}},$$

  et cela reste valable pour $x = 1$ par continuité de g'.

- Valeur de g

  Ainsi : $\forall\, x > 0,\ g'(x) = \dfrac{\pi}{2}\dfrac{1}{x+\sqrt{x}}$ *et $g(1) = 0$ (calcul immédiat), d'où :*

  $$\forall\, x > 0,\ g(x) = \frac{\pi}{2}\int_1^x \frac{\mathrm{d}t}{t+\sqrt{t}} = \pi\int_1^x \frac{\mathrm{d}t}{2\sqrt{t}(1+\sqrt{t})} = \pi\Big[\ln(\sqrt{t}+1)\Big]_1^x$$

  soit enfin : $\forall\, x > 0,\ g(x) = \pi\ln\left(\dfrac{\sqrt{x}+1}{2}\right)$.

- Continuité de g

  Nous appliquons le théorème sur la continuité d'une intégrale à paramètre.

  – *La fonction $(x,t) \mapsto f(x,t) = \ln(x\cos^2 t + \sin^2 t)$ est continue sur $\mathbb{R}_+ \times \left]0\,;\frac{\pi}{2}\right]$ (vous remarquerez que, puisque ici x peut prendre la valeur 0, on applique le théorème pour t dans l'intervalle semi-ouvert $\left]0\,;\frac{\pi}{2}\right]$).*

– *Il ne reste plus qu'à vérifier l'hypothèse de domination.*
On peut se contenter de démontrer la continuité de g sur $[0\,;1]$ (la continuité sur $\mathbb{R}_+^*$ découlant de sa dérivabilité).
Supposons donc $x \in [0\,;1]$. On a alors l'encadrement :

$$\forall\, t \in \left]0\,;\tfrac{\pi}{2}\right],\ \sin^2 t \leqslant x\cos^2 t + \sin^2 t \leqslant 1\,,$$

d'où :

$$\forall\, x \in [0\,;1],\ \forall\, t \in \left]0\,;\tfrac{\pi}{2}\right],\ |f(x,t)| \leqslant \left|\ln(\sin^2 t)\right| = 2\,|\ln\sin t| = \varphi(t)\,,$$

où φ est continue et intégrable sur $\left]0\,;\frac{\pi}{2}\right]$ (vérification facile, puisque $\ln\sin t \underset{t\to 0^+}{\sim} \ln t$). L'hypothèse de domination est donc vérifiée.

On en déduit que g est continue sur $[0\,;1]$, donc sur $\mathbb{R}_+$.

- *g étant continue en 0^+ : $g(0) = \lim\limits_{x\to 0^+} g(x) = \displaystyle\int_0^{\frac{\pi}{2}} \ln(\sin^2 t)\,\mathrm{d}t$.*
Le calcul de $g(x)$ pour $x > 0$ fait précédemment montre que cette limite est égale à $-\pi\ln 2$. On en déduit l'égalité : $\displaystyle\int_0^{\frac{\pi}{2}} \ln(\sin t)\,\mathrm{d}t = -\frac{\pi}{2}\ln 2$.

Exemple

Calculer pour $x \geqslant 0$: $g(x) = \displaystyle\int_0^{+\infty} \frac{\sin(xt)}{t(1+t^2)}\,\mathrm{d}t$ (on rappelle : $\displaystyle\int_0^{+\infty} \frac{\sin u}{u}\,\mathrm{d}u = \frac{\pi}{2}$).

Posons pour tout $(x,t) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, $f(x,t) = \dfrac{\sin(xt)}{t(1+t^2)}$.

- *Montrons d'abord que g est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$; pour cela on applique directement le théorème du cours.*
  – *Pour tout $x \in \mathbb{R}_+$, $t \mapsto f(x,t)$ est continue sur $\mathbb{R}_+^*$ d'après les théorèmes usuels. Elle y est de plus intégrable car $\lim\limits_{t\to 0^+} f(x,t) = x$ et $|f(x,t)| \leqslant \dfrac{1}{t^3}$ qui est une fonction intégrable au voisinage de $+\infty$.*
  – *f admet en tout point $(x,t) \in \mathbb{R}_+ \times \mathbb{R}_+^*$ une dérivée partielle $\dfrac{\partial f}{\partial x}(x,t) = \dfrac{\cos(xt)}{1+t^2}$, qui est évidemment continue par rapport à x et par rapport à t.*
  – $\forall\,(x,t) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, $\left|\dfrac{\partial f}{\partial x}(x,t)\right| = \left|\dfrac{\cos(xt)}{1+t^2}\right| \leqslant \dfrac{1}{1+t^2}$, *avec* $t \mapsto \dfrac{1}{1+t^2}$ *continue et intégrable sur $\mathbb{R}_+^*$ (facile).*
  D'après le théorème relatif à la dérivation d'une intégrale dépendant d'un paramètre, la fonction g est donc définie et de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$, et

$$\forall\, x \geqslant 0,\ g'(x) = \int_0^{+\infty} \frac{\cos(xt)}{1+t^2}\,\mathrm{d}t\,.$$

- *On ne peut pas calculer cette intégrale, on va donc plutôt chercher une équation différentielle vérifiée par g. Pour cela, il faudrait pouvoir dériver une fois de plus afin de faire apparaître un $\sin$; mais malheureusement l'expression de $g'(x)$ ci-dessus ne permet pas d'appliquer de nouveau le théorème de dérivation d'une intégrale à paramètre (essayez !).*

On va donc utiliser une méthode déjà mentionnée en **[SF14.4]** *: faire une intégration par parties afin d'élever le degré du dénominateur.*
Pour tout $x > 0$ *on a ainsi :*

$$g'(x) = \left[\frac{\sin xt}{x} \cdot \frac{1}{1+t^2}\right]_{t=0}^{t=+\infty} + \int_0^{+\infty} \frac{\sin xt}{x} \cdot \frac{2t}{(1+t^2)^2}\,\mathrm{d}t$$

où l'intégration par parties est justifiée car le terme entre crochets admet une limite, d'ailleurs nulle, en $+\infty$.

Ainsi pour $x > 0$ $g'(x) = \dfrac{2}{x}\displaystyle\int_0^{+\infty} \frac{t\sin xt}{(1+t^2)^2}\,\mathrm{d}t$. *Sous cette forme on peut appliquer à nouveau le théorème de dérivation d'une intégrale à paramètre (les hypothèses sont faciles à vérifier, nous ne détaillons pas les calculs), et l'on obtient pour* $x > 0$ *l'expression suivante.*

$$\begin{aligned}
g''(x) &= -\frac{2}{x^2}\int_0^{+\infty} \frac{t\sin xt}{(1+t^2)^2}\,\mathrm{d}t + \frac{2}{x}\int_0^{+\infty} \frac{t^2\cos xt}{(1+t^2)^2}\,\mathrm{d}t \\
&= -\frac{1}{x}g'(x) + \frac{1}{x}\int_0^{+\infty} \left(\frac{2t}{(1+t^2)^2}\right)\cdot t\cos xt\,\mathrm{d}t \\
&= -\frac{1}{x}g'(x) + \frac{1}{x}\left[-\frac{1}{1+t^2}\cdot t\cos xt\right]_{t=0}^{t=+\infty} \\
&\qquad + \frac{1}{x}\int_0^{+\infty} \frac{1}{1+t^2}(\cos xt - xt\sin xt)\,\mathrm{d}t \\
&= \underbrace{-\frac{1}{x}g'(x) + \frac{1}{x}\int_0^{+\infty} \frac{\cos xt}{1+t^2}\,\mathrm{d}t}_{=0} - \int_0^{+\infty} \frac{t\sin xt}{1+t^2}\,\mathrm{d}t\,.
\end{aligned}$$

On en déduit pour $x > 0$ *:*

$$g''(x) - g(x) = \int_0^{+\infty} \frac{\sin xt}{1+t^2}\left(-t - \frac{1}{t}\right)\mathrm{d}t = -\int_0^{+\infty} \frac{\sin xt}{t}\,\mathrm{d}t\,.$$

En effectuant le changement de variable affine $u = xt$*, on obtient :*

$$g''(x) - g(x) = -\int_0^{+\infty} \frac{\sin u}{u}\,\mathrm{d}u = -\frac{\pi}{2}\cdot$$

On en déduit pour $x > 0$ $g(x) = \dfrac{\pi}{2} + \lambda\mathrm{e}^{x} + \mu\mathrm{e}^{-x}$*, avec* $(\lambda,\mu) \in \mathbb{R}^2$*, et par continuité de* g *cette formule reste valable pour* $x = 0$.
Comme $g(0) = 0$ *et* $g'(0) = \dfrac{\pi}{2}$ *on en déduit finalement :*

$$\forall\, x \geqslant 0,\ g(x) = \frac{\pi}{2}(1 - \mathrm{e}^{-x})\,.$$

■ Comment intervertir série-intégrale ?

SF14.23 Utiliser un théorème du cours

Il y a trois théorèmes dans le cours qui permettent de justifier l'interversion série-intégrale.

– Le premier, rappelé en **[S12.10]**, s'applique exclusivement lorsque l'intervalle d'intégration est un segment et lorsque la convergence de la série est uniforme.
Nous avons déjà donné des exemples d'application de ce théorème en **[SF12.20]**.

– Les deux autres sont rappelés en **[S14.10]** : il s'agit du théorème d'intégration terme à terme et du théorème de convergence dominée appliqué aux séries de fonctions.
Pour savoir lequel employer, on peut utiliser les remarques ci-dessous (mais attention, il peut y avoir des exceptions !).

  – Si les fonctions u_n sont *à valeurs réelles et de signe constant* sur l'intervalle d'intégration I, essayez d'appliquer en premier le théorème d'intégration terme à terme. Le seul problème de ce théorème est qu'il faut savoir calculer $\int_I |u_n|$, mais cela revient à calculer $\int_I u_n$ puisque u_n est de signe constant, et il faudra le faire de toutes façons si on veut pouvoir écrire $\displaystyle\int_I \left(\sum_{n=0}^{+\infty} u_n\right) = \sum_{n=0}^{+\infty} \int_I u_n$.
  Dans ce cas, c'est ce théorème qui comporte alors le moins d'hypothèses.

  – Le théorème de convergence dominée est plutôt à utiliser dans le cas de fonctions de signes non constants, ou à valeurs dans $\mathbb{C}$. En effet dans ce cas, il se peut que le calcul de $\displaystyle\int_I |u_n|$ soit plus délicat que celui de $\displaystyle\int_I u_n$, ou même que la série $\sum \displaystyle\int_I |u_n|$ soit divergente alors que la série $\sum \displaystyle\int_I u_n$ est convergente.
  Le problème de ce théorème est qu'il faut arriver à majorer les modules ou valeurs absolues des sommes partielles par une fonction intégrable, ce qui n'est pas toujours facile.

Exemple

Montrer que : $\displaystyle\int_0^1 \frac{\ln x}{1+x^2}\,\mathrm{d}x = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{(2n+1)^2}$.

La fonction $f : x \mapsto \dfrac{\ln x}{1+x^2}$ *est continue sur* $]0\,;1]$, *et* $f(x) \underset{x\to 0}{\sim} \ln x$ *qui est une fonction intégrable, l'intégrale proposée existe donc.*
D'après les développements en série entière usuels, on a :

$$\forall\, x \in]0\,;1[,\ \frac{1}{1+x^2} = \sum_{k=0}^{+\infty} (-1)^k x^{2k}\,,$$

donc, pour $x \in [0\,;1[$, $\displaystyle f(x) = \sum_{k=0}^{+\infty} (-1)^k x^{2k} \ln x$.
Posons $u_k(x) = (-1)^k x^{2k} \ln x$. *Vérifions les hypothèses du théorème d'intégration terme à terme.*

– *Les fonctions* u_k *sont continues et intégrables sur* $]0\,;1[$ *(pour* $k = 0$, *il s'agit de la fonction* ln *qui est une fonction de référence, et pour* $k \geqslant 1$, u_k *est prolongeable en* 0*).*

– *D'après le calcul précédent, la série de fonctions* $\sum u_k$ *converge simplement sur* $]0\,;1[$ *vers* f.

– *Enfin,* $\displaystyle\int_0^1 |u_k| = -\int_0^1 x^{2k}\ln x\,\mathrm{d}x = \frac{1}{(2k+1)^2}$ *en faisant une intégration par parties, donc la série* $\sum\int|u_k|$ *est convergente.*

On peut donc appliquer le théorème d'intégration terme à terme qui donne :

$$\int_0^1 f(x)\,\mathrm{d}x = \sum_{k=0}^{+\infty}\int_0^1 u_k(x)\,\mathrm{d}x = \sum_{k=0}^{+\infty}\frac{(-1)^{k+1}}{(2k+1)^2}\,.$$

✓ *Le nombre trouvé s'appelle la* constante de Catalan. *On ne sait pas si ce nombre est rationnel ou irrationnel.*

Exemple

Montrer que : $\displaystyle\int_0^{+\infty}\frac{\sin x\,\mathrm{d}x}{\mathrm{e}^x-1} = \sum_{n=1}^{+\infty}\frac{1}{n^2+1}\,.$

- *La fonction* $f\colon x\mapsto\dfrac{\sin x}{\mathrm{e}^x-1}$ *est continue sur* $\mathbb{R}_+^*$, *et prolongeable par continuité en* 0, *puisque* $\lim\limits_{x\to0^+} f(x)=1$ *(car* $\sin x\underset{0}{\sim}x$ *et* $\mathrm{e}^x-1\underset{0}{\sim}x$*).*
  Elle est intégrable sur $\mathbb{R}_+$, *puisque* $\lim\limits_{x\to+\infty}x^2f(x)=0$ *donc* $f(x)\underset{+\infty}{=}o\left(\dfrac{1}{x^2}\right)$ *(comparaison à une intégrale de Riemann).*

- *Le développement en série entière de* $\dfrac{1}{1-X}$ *pour* $|X|<1$ *conduit à :*

$$\forall\,x>0,\ f(x)=\frac{\mathrm{e}^{-x}\sin x}{1-\mathrm{e}^{-x}}=\sum_{n=0}^{+\infty}\sin x.\mathrm{e}^{-(n+1)x}\,.$$

Pour $n\in\mathbb{N}$ *et* $x\in\mathbb{R}_+$, *notons* $u_n(x)=\sin x.\mathrm{e}^{-(n+1)x}$. *Alors* u_n *est continue et intégrable sur* $\mathbb{R}_+$ *(en majorant* $|u_n(x)|\leqslant\mathrm{e}^{-(n+1)x}$*), et la série de fonctions* $\sum u_n$ *converge simplement sur* $\mathbb{R}_+^*$ *vers* f.
En posant, pour $N\in\mathbb{N}$, $f_N(x)=\sum\limits_{n=0}^{N}u_n(x)$, *les fonctions* f_N *sont donc continues et intégrables sur* $\mathbb{R}_+$, *et la suite* (f_N) *convergent simplement sur* $\mathbb{R}_+$ *vers* f. *De plus, on a :*

$$|f_N(x)|\leqslant\sum_{n=0}^{+\infty}|\sin x|\mathrm{e}^{-(n+1)x}=\frac{|\sin x|}{\mathrm{e}^x-1}=\varphi(x)\,,$$

avec φ *continue et intégrable sur* $]0\,;+\infty[$ *(vérification facile).*
Le théorème de convergence dominée appliqué à la suite (f_N) *donne alors :*

$$\begin{aligned}\int_0^{+\infty}f(x)\,\mathrm{d}x &= \lim_{N\to+\infty}\int_0^{+\infty}f_N(x)\,\mathrm{d}x\\ &= \lim_{N\to+\infty}\int_0^{+\infty}\sum_{n=0}^{N}u_n(x)=\lim_{N\to+\infty}\sum_{n=0}^{N}\int_0^{+\infty}u_n(x)\,\mathrm{d}x\,.\end{aligned}$$

On calcule :
$$\int_0^{+\infty} u_n(x)\,\mathrm{d}x = \mathcal{I}m\left(\int_0^{+\infty} \mathrm{e}^{[\mathrm{i}-(n+1)]x}\,\mathrm{d}x\right)$$
$$= \mathcal{I}m\left(\left[\frac{1}{\mathrm{i}-(n+1)}\mathrm{e}^{[\mathrm{i}-(n+1)]x}\right]_0^{+\infty}\right)$$
$$= \mathcal{I}m\left(\frac{1}{n+1-\mathrm{i}}\right) = \frac{1}{(n+1)^2+1}$$

d'où finalement : $\displaystyle\int_0^{+\infty} f(x)\,\mathrm{d}x = \sum_{n=0}^{+\infty} \frac{1}{(n+1)^2+1} = \sum_{n=1}^{+\infty} \frac{1}{n^2+1}\,.$

✓ *Il ne semble pas possible dans cet exemple d'appliquer le théorème d'intégration terme à terme. En effet, on a seulement :*

$$\int_0^{+\infty} |u_n(x)|\,\mathrm{d}x \leqslant \int_0^{+\infty} \mathrm{e}^{-(n+1)x}\,\mathrm{d}x = \frac{1}{n+1}\,,$$

et la série de terme général $\dfrac{1}{n+1}$ *ne converge pas...*

SF14.24 Faire une démonstration directe

Dans le cas où l'on ne peut pas appliquer l'un des théorèmes précédents (ou si l'on n'arrive pas à en vérifier les hypothèses), on revient simplement à la définition d'une série en considérant les sommes partielles et les restes.

Nous avons expliqué cette méthode en **[SF12.21]** page 500. Elle consite à démontrer que l'intégrale du reste de la série tend vers 0. Voyez l'exemple de **[SF12.21]** page 500, et aussi l'exercice **14.9** ci-après.

À vous de jouer !

Exercice 14.1 **SF14.3**

1) Justifier l'intégrabilité sur $]1\,;+\infty[$ de $f\colon x \mapsto \dfrac{\sqrt{\ln x}}{(x-1)\sqrt{x}}\,.$

2) Justifier l'intégrabilité sur $]0\,;1[$ de $g\colon x \mapsto \dfrac{\sqrt{|\ln x|}}{(x-1)\sqrt{x}}\,.$

Exercice 14.2 **SF14.1** **SF14.6**

Justifier l'existence de l'intégrale : $\displaystyle\int_0^{+\infty} \ln(\mathrm{th}\,t)\,\mathrm{d}t\,.$

Exercice 14.3 **SF14.7**

Étudier l'existence de l'intégrale $\displaystyle\int_0^{+\infty} t\mathrm{e}^{-t^3|\sin t|}\,\mathrm{d}t$.

Exercice 14.4 **SF14.8** **SF14.10**

Soit f continue sur $\mathbb{R}_+$. On définit une fonction g sur $\mathbb{R}_+$ par :

$$\forall\, x>0,\ g(x)=\frac{1}{x}\int_0^x f(t)\,\mathrm{d}t \quad \text{et} \quad g(0)=f(0)\,.$$

1) Montrer que g est continue sur $\mathbb{R}_+$.

2) On suppose maintenant que $\displaystyle\int_0^{+\infty} f^2(t)\,\mathrm{d}t$ est convergente.

a) On se donne a et b deux réels tels que $0<a<b$.
Montrer que : $\displaystyle\int_a^b g^2(t)\,\mathrm{d}t = 2\int_a^b f(t)g(t)\,\mathrm{d}t + ag^2(a) - bg^2(b)$.

b) En déduire que l'intégrale $\displaystyle\int_0^{+\infty} g^2(t)\,\mathrm{d}t$ est convergente et que :

$$\int_0^{+\infty} g^2(t)\,\mathrm{d}t \leqslant 4\int_0^{+\infty} f^2(t)\,\mathrm{d}t\,.$$

c) Démontrer que $\displaystyle\lim_{x\to+\infty} xg^2(x)=0$.

Exercice 14.5 **SF14.9** **SF14.21**

Soit f continue de $[1\,;+\infty]$ dans $\mathbb{R}$. Montrer si l'une des deux limites suivantes existe il en est de même de l'autre, et qu'elles sont alors égales :

$$\lim_{x\to+\infty}\frac{1}{x}\int_1^x f(t)\,\mathrm{d}t \quad \text{et} \quad \lim_{x\to+\infty} x\int_x^{+\infty}\frac{f(t)}{t^2}\,\mathrm{d}t\,.$$

Exercice 14.6 **SF14.8** **SF14.9**

Soit $\alpha>0$, et $f\colon [1\,;+\infty[\to\mathbb{R}$ continue telle que $\displaystyle\int_1^{+\infty} f$ existe.

Montrer que $\displaystyle\int_1^{+\infty}\frac{f(t)}{t^\alpha}\,\mathrm{d}x$ existe.

Exercice 14.7 **SF14.3** **SF14.21**

Soit f de classe $\mathscr{C}^1$ de $\mathbb{R}_+$ dans $\mathbb{R}_+^*$; on suppose que $\alpha=\displaystyle\lim_{x\to+\infty}\frac{xf'(x)}{f(x)}$ existe.

1) Montrer que $\displaystyle\lim_{x\to+\infty}\frac{\ln f(x)}{\ln x}=\alpha$.

2) Si $\alpha < -1$, montrer que f est intégrable sur $\mathbb{R}_+$ et que :

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} -\frac{xf(x)}{\alpha+1}\,.$$

3) Si $\alpha > -1$, montrer que f n'est pas intégrable sur $\mathbb{R}_+$ et que :

$$\int_0^{x} f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} \frac{xf(x)}{\alpha+1}\,.$$

Exercice 14.8 **SF10.13** **SF14.3** **SF14.22** **SF14.23**

1) Montrer que f, définie par $f(x) = \displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-t^2}\,\mathrm{d}t$, est de classe $\mathscr{C}^1$ sur $\mathbb{R}$.

2) **a)** Déterminer une équation différentielle linéaire du 1er ordre vérifiée par f.

b) En admettant que $f(0) = \sqrt{\pi}$, calculer $f(x)$ pour tout x.

3) Démontrer que la fonction $g\colon x \mapsto \displaystyle\int_0^{+\infty} \mathrm{e}^{-xt}\sin\sqrt{t}\,\mathrm{d}t$ est bien définie pour tout $x > 0$.
Trouver une relation entre f et g, et en déduire la valeur de g.

4) Calculer directement g par une autre méthode.

Exercice 14.9 **SF12.12** **SF14.22** **SF14.23** **SF14.24**

Pour $x \in \mathbb{R}$ on pose $F(x) = \displaystyle\int_0^{+\infty} \frac{\sin tx}{\mathrm{e}^t - 1}\,\mathrm{d}t\,.$

1) Montrer que F est continue sur $\mathbb{R}$.

2) Montrer que pour tout $x \in \mathbb{R}$, $F(x) = \displaystyle\sum_{n=1}^{+\infty} \frac{x}{x^2+n^2}\,.$

3) En déduire $\displaystyle\lim_{x\to+\infty} F(x)$.

Exercice 14.10 **SF14.3** **SF14.12** **SF14.23**

Pour $n \in \mathbb{N}^*$, on pose $u_n = \displaystyle\int_0^{+\infty} \frac{\mathrm{d}x}{\mathrm{ch}^n x}$.

1) Montrer que u_n existe pour tout $n \in \mathbb{N}^*$.

2) Montrer que la suite (u_n) converge vers une limite que l'on précisera.

3) **a)** Déterminer la nature et la somme éventuelle de la série $\displaystyle\sum_{n\in\mathbb{N}^*} u_n$.

b) Déterminer la nature et la somme éventuelle de la série $\displaystyle\sum_{n\in\mathbb{N}^*} (-1)^n u_n$.

Thème 15 - Espaces probabilisés

■ Comment définir une probabilité ?

SF15.1 Revenir à la définition axiomatique

On peut en théorie définir une probabilité sur un ensemble Ω en construisant une tribu $\mathcal{A}$ de parties de Ω et une application P de $\mathcal{A}$ dans $[0\,;1]$ qui vérifient les axiomes rappelés en **[S15.3]**.

C'est de cette façon qu'il a été démontré en cours qu'une probabilité conditionnelle définit bien une probabilité.

Mais hormis ce cas, cela n'est jamais simple, et ne présente de véritable intérêt que lorsque l'univers Ω n'est pas dénombrable, ce qui sort du cadre du programme.

SF15.2 Utiliser un germe de probabilité

Il s'agit de la méthode qui est la plus utilisée dans le cas où l'univers Ω est au plus dénombrable. En effet dans ce cas, pour définir une probabilité, il suffit de la définir pour tous les singletons $\{\omega\}$, de s'assurer que les réels $P(\{\omega\})$ sont positifs et de vérifier que $\sum_{\omega\in\Omega} P(\{\omega\}) = 1$.

Exemple

1) Montrer que l'on peut définir sur $\mathbb{N}$ une probabilité par :

$$\forall\, n \in \mathbb{N},\ P(\{n\}) = \frac{1}{2^{n+1}}\cdot$$

2) Soit $k \in \mathbb{N}$. Quelle est la probabilité qu'un entier soit multiple de k ?

1) *Les nombres $P(\{n\})$ étant positifs, pour répondre à la question il suffit de remarquer que :*

$$\sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}} = 1\,,$$

d'après les formules sur les séries géométriques.

2) *La probabilité cherchée est $P(k\mathbb{N})$. On a directement, par σ-additivité :*

$$P(k\mathbb{N}) = \sum_{x\in k\mathbb{N}} P(\{x\}) = \sum_{n=0}^{+\infty} P(\{kn\}) = \sum_{n=0}^{+\infty} \frac{1}{2^{kn+1}} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{2^k}} = \frac{2^{k-1}}{2^k-1}\cdot$$

✓ *Pour cette loi, la probabilité qu'un entier soit pair vaut : $P(2\mathbb{N}) = \dfrac{2}{3}\cdot$*

Cela ne correspond donc pas à la notion intuitive que l'on peut avoir...

■ Comment calculer la probabilité d'un évènement ?

Les méthodes de calcul d'une probabilité sont nombreuses, et la difficulté vient souvent du fait que plusieurs d'entre elles doivent être employées consécutivement.

Il est donc important d'essayer de découper le problème posé en différents problèmes plus simples, ceux-ci correspondant souvent aux différentes étapes de l'expérience aléatoire étudiée.

Pour cela il est *indispensable* de donner un nom aux évènements considérés par l'énoncé et à ceux qui le composent, afin d'écrire l'évènement étudié comme combinaison (réunion, intersection, complémentaire...) d'évènements plus simples.

SF15.3 Utiliser la formule de la probabilité uniforme

Dans le cas où l'univers Ω est de cardinal fini, on le munit généralement de la probabilité uniforme, qui est telle que chaque singleton a la même probabilité.
Dans ce cas, le calcul de la probabilité d'un évènement A se ramène à un « simple » problème de dénombrement, puisque l'on a alors :

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)},$$

ce que l'on résume en : $P(A) = \dfrac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$.

✓ On fera cependant attention à ne pas utiliser cette formule si l'hypothèse d'équiprobabilité des évènements élémentaires n'est pas vérifiée.

Ce type de situation est important, et c'est principalement celui que vous avez étudié en 1re année. Nous nous contentons donc de le rappeler sans donner d'exemple particulier, mais nous aurons plusieurs fois l'occasion de l'utiliser.

SF15.4 Passer par l'évènement contraire

Si A est un évènement, il est parfois plus facile de calculer la probabilité de l'évènement contraire $\overline{A}$. Les deux probabilités sont reliées par la relation :

$$P(\overline{A}) = 1 - P(A).$$

On emploie généralement cette méthode lorsque la description de l'évènement comporte la locution « au moins », l'évènement contraire de « A survient au moins une fois » étant « A ne survient jamais ».

Exemple

Le nombre d'erreurs par page de ce livre est une variable aléatoire suivant une loi de Poisson de paramètre λ (λ petit...)

Quelle est la probabilité qu'il y ait au moins une erreur sur cette page ?

Si l'on note X le nombre d'erreurs sur cette page, on cherche $P(X \geqslant 1)$.
En utilisant l'évènement contraire :

$$P(X \geqslant 1) = 1 - P(X = 0) = 1 - \mathrm{e}^{-\lambda}.$$

Exemple

N personnes sont réunies. Quelle est la probabilité pour qu'au moins 2 d'entre elles aient le même jour anniversaire ? (On supposera qu'aucune n'est née un 29 février, et que $N \leqslant 365$).

L'évènement contraire de l'évènement :

A : « au moins 2 personnes ont le même jour anniversaire »

est l'évènement :

$\overline{A}$: « aucune des N personnes n'a le même jour anniversaire ».

On peut prendre ici comme univers Ω l'ensemble $[\![1\,;365]\!]^N$ formé des N-uplets comportant les jours anniversaires possibles de chaque personne. L'évènement $\overline{A}$ est alors l'ensemble des N-listes d'entiers distincts entre 1 et 365 (ou N-arrangements, voir **[S15.1]***), donc :*

$$\mathrm{Card}(\overline{A}) = 365 \times 364 \times \cdots \times (365 - N + 1) = \frac{365!}{(365-N)!}\,.$$

La probabilité cherchée vaut donc :

$$P(A) = 1 - \frac{365 \times 364 \times \cdots \times (365 - N + 1)}{365^N}\,.$$

Pour l'anecdote, notons que $P(A) > 0{,}5$ dès que $N \geqslant 23$ et que $P(A) > 0{,}99$ dès que $N \geqslant 57$.

SF15.5 Écrire l'évènement comme réunion de deux évènements

Il s'agit ici d'utiliser la relation :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,.$$

En particulier si A et B sont incompatibles, on aura $P(A \cup B) = P(A) + P(B)$.

Exemple

On dispose de 5 paires de chaussettes que l'on range au hasard dans 3 tiroirs (chaque tiroir peut contenir plusieurs paires).

Quelle est la probabilité qu'un tiroir au moins soit vide ?

Il y a 3^5 façons de répartir les paires de chaussettes (pour chaque paire, il y a 3 tiroirs possibles). Considérons pour $k \in \{1,2,3\}$ les évènements :

A_k : « le tiroir n^o k est vide ».

L'évènement considéré est donc $B = A_1 \cup A_2 \cup A_3$. Donc :

$$\begin{aligned} P(B) &= P(A_1 \cup A_2) + P(A_3) - P\big((A_1 \cup A_2) \cap A_3\big) \\ &= P(A_1) + P(A_2) - P(A_1 \cap A_2) + P(A_3) - P\big((A_1 \cap A_3) \cup (A_2 \cap A_3)\big) \\ &= \sum_{i=1}^{3} P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + P(A_1 \cap A_2 \cap A_3)\,. \end{aligned}$$

Calculons les probabilités de ces différents évènements :

- *pour tout $i \in \{1,2,3\}$, $P(A_i) = \dfrac{2^5}{3^5}$ puisqu'il y a 2^5 façons de répartir les 5 paires dans les deux autres tiroirs ;*

– *pour* $i \neq j$, $P(A_i \cap A_j) = \dfrac{1}{3^5}$ *: il n'y a alors qu'une disposition possible ;*
– *et enfin* $P(A_1 \cap A_2 \cap A_3) = 0$, *la configuration imposée étant impossible.*

On en déduit : $P(B) = 3 \cdot \dfrac{2^5}{3^5} - 3 \cdot \dfrac{1}{3^5} = \dfrac{31}{81}$.

✓ *Chercher l'évènement contraire de* B *revient à chercher le nombre d'applications surjectives d'un ensemble à 5 éléments dans un ensemble à 3 éléments ; or calculer le nombre d'applications surjectives d'un ensemble fini dans un autre est possible, mais pas simple. Pour cette raison, contrairement à la plupart des évènements dont l'énoncé comporte la locution « au moins », nous n'avons pas utilisé ici l'évènement contraire.*

SF15.6 Écrire l'évènement comme réunion finie d'évènements

Il s'agit ici d'écrire un évènement A sous la forme : $A = \bigcup_{k=1}^{n} A_k$.

- Si les évènements A_k sont deux à deux incompatibles, on aura alors :
$$P(A) = \sum_{k=1}^{n} P(A_k).$$
- Sinon, on peut essayer de passer aux évènements contraires en utilisant :
$$P(\overline{A}) = P\left(\overline{\bigcup_{k=1}^{n} A_k}\right) = P\left(\bigcap_{k=1}^{n} \overline{A_k}\right).$$

Exemple

On permute au hasard les tomes d'une encyclopédie, numérotés de 1 à n.

Quelle est la probabilité que les tomes 1 et 2 se retrouvent côte à côte dans cet ordre ?

L'univers Ω *est ici l'ensemble des permutations de* $[\![1\,;n]\!]$, *il est donc de cardinal* $n!$.
Si l'on note A_k, *pour* $k \in [\![1\,;n-1]\!]$, *l'évènement :*

« le tome 1 *est à la* k^e *place et le tome* 2 *à la* $(k+1)^e$ *place »,*

et A *l'évènement :*

« les tomes 1 *et* 2 *se retrouvent côte à côte dans cet ordre »,*

l'évènement A *est la réunion disjointe des évènements* A_k.
On a donc $P(A) = \sum_{k=1}^{n-1} P(A_k)$.
Or le cardinal de chaque A_k *est égal à* $(n-2)!$ *puisque si les tomes* 1 *et* 2 *sont placés, il reste* $(n-2)!$ *façons de disposer les autres dans les* $n-2$ *places restantes.*
Donc $P(A_k) = \dfrac{(n-2)!}{n!}$ *puis* $P(A) = (n-1)\dfrac{(n-2)!}{n!} = \dfrac{1}{n}$.

Exemple

Soient $A_1, \ldots, A_n$ n évènements d'un espace probabilisé $(\Omega, \mathcal{A}, P)$, mutuellement indépendants.

Exprimer $P(A_1 \cup \ldots \cup A_n)$ en fonction des $P(A_i)$.

Si les A_i sont mutuellement indépendants, il en est de même des évènements contraires $\overline{A_i}$. On a donc :

$$P(A_1 \cup \ldots \cup A_n) = 1 - P(\overline{A_1} \cap \ldots \cap \overline{A_n}) = 1 - \prod_{i=1}^{n} P(\overline{A_i}) = 1 - \prod_{i=1}^{n} (1 - P(A_i)) .$$

SF15.7 Écrire l'évènement comme réunion infinie d'évènements

Il s'agit ici d'écrire un évènement A sous la forme : $A = \bigcup_{n=1}^{+\infty} A_n$.

- Si les évènements A_n sont deux à deux incompatibles, on aura :
$$P(A) = \sum_{n=1}^{+\infty} P(A_n) .$$
- Si la suite d'évènements (A_n) est croissante, c'est-à-dire $A_n \subset A_{n+1}$, on aura :
$$P(A) = \lim_{n \to +\infty} P(A_n) .$$
- On peut aussi utiliser la formule : $P(A) = \lim_{n \to +\infty} P\left(\bigcup_{k=1}^{n} A_k \right)$.
- Enfin, on peut aussi essayer de calculer la probabilité de l'évènement contraire :
$$\overline{\bigcup_{n=1}^{+\infty} A_n} = \bigcap_{n=1}^{+\infty} \overline{A_n} .$$

Exemple

On lance un dé équilibré jusqu'à l'obtention d'un 6.

Quelle est la probabilité que tous les nombres obtenus soient pairs ?

Pour tout entier $n \geqslant 1$, notons A_n l'évènement :

« les $(n-1)^{ers}$ lancers donnent 2 ou 4 et le n^e donne 6 ».

L'évènement considéré par l'énoncé est $A = \bigcup_{n=1}^{+\infty} A_n$.

Les A_n étant disjoints, on a donc $P(A) = \sum_{n=1}^{+\infty} P(A_n)$.

Or $P(A_n) = \left(\frac{1}{3}\right)^{n-1} \times \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3^n}$, donc la probabilité cherchée est :

$$P(A) = \frac{1}{2} \sum_{n=1}^{+\infty} \frac{1}{3^n} = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{1 - \frac{1}{3}} = \frac{1}{4} .$$

Exemple

On lance indéfiniment une pièce équilibrée.

Quelle est la probabilité d'apparition de la séquence (pile, pile, face) ?

Considérons les évènements :

– P_n : « *on obtient pile au n^e lancer* » $(n \geqslant 1)$;
– F_n : « *on obtient face au n^e lancer* » $(n \geqslant 1)$;
– $B_n = P_{n-2} \cap P_{n-1} \cap F_n$ $(n \geqslant 3)$;
– $A_n = \bigcup_{k=3}^{n} B_k$ $(n \geqslant 3)$.

Ainsi, A_n correspond à l'évènement :

« la séquence (pile, pile, face) est apparue avant le n^e lancer »,

et l'évènement considéré par l'énoncé :

A : « on obtient au moins une fois la séquence (pile, pile, face) »

est la réunion des A_n.

Puisque $A_{n+1} = A_n \cup B_{n+1}$ on a :

$$P(A_{n+1}) = P(A_n) + P(B_{n+1}) - P(A_n \cap B_{n+1}) \quad (*)$$

Mais pour $n \geqslant 5$:

$$A_n \cap B_{n+1} = (A_{n-2} \cap B_{n+1}) \cup \underbrace{(B_{n-1} \cap B_{n+1})}_{=\emptyset} \cup \underbrace{(B_n \cap B_{n+1})}_{=\emptyset}$$
$$= A_{n-2} \cap B_{n+1}\,,$$

donc $P(A_n \cap B_{n+1}) = P(A_{n-2} \cap B_{n+1}) = P(A_{n-2})P(B_{n+1})$ puisque les évènements A_{n-2} et B_{n+1} sont indépendants (ce dernier ne faisant intervenir que les tirages après le $(n-1)^e$).

Puisque clairement $P(B_n) = \dfrac{1}{8}$ pour tout $n \in \mathbb{N}^*$, la relation $(*)$ conduit finalement à la relation de récurrence (pour $n \geqslant 5$) :

$$P(A_{n+1}) = P(A_n) + \frac{1}{8} - \frac{1}{8}P(A_{n-2}) = P(A_n) + \frac{1}{8}\big(1 - P(A_{n-2})\big)\,.$$

Cela implique en particulier que la suite $(P(A_n))_{n\geqslant 5}$ est croissante ; étant majorée par 1 elle converge et par passage à la limite dans la relation de récurrence, sa limite ℓ vérifie :

$$\ell = \ell + \frac{1}{8}(1-\ell)\,, \quad \textit{soit} \quad \ell = 1\,.$$

Pour terminer, on remarque que la suite (A_n) est croissante, donc

$$P(A) = P\left(\bigcup_{n=3}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) = 1\,.$$

L'évènement « on obtient au moins une fois la séquence (pile, pile, face) » est donc quasi-certain.

SF15.8 Écrire l'évènement comme intersection finie d'évènements

Il s'agit ici d'écrire un évènement A sous la forme : $A = \bigcap_{k=1}^{n} A_k$.

- Si les évènements A_k sont mutuellement indépendants, on aura :

$$P(A) = \prod_{k=1}^{n} P(A_k)\,.$$

- Dans le cas général, on utilisera la formule des probabilités composées :

$$P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1)P\left(A_2|A_1\right)P\left(A_3|A_1\cap A_2\right)\ldots P\left(A_n|A_1\cap A_2\cap\ldots\cap A_{n-1}\right)$$

  sous réserve que $P(A_1\cap A_2\cap\ldots\cap A_{n-1})\neq 0$.

- On peut aussi essayer de passer à l'évènement contraire, ce qui nous ramène au calcul de la probabilité d'une réunion finie d'évènements.

Exemple

Un livre contient N erreurs ($N\in\mathbb{N}^*$) et est relu par une suite de correcteurs. On suppose qu'à chaque relecture, chaque erreur est corrigée avec la probabilité p. Les erreurs sont corrigées de manière indépendante les unes des autres, et les relectures sont indépendantes les unes des autres.

Quelle est la probabilité que le livre soit entièrement corrigé après n relectures ?

Pour $i\in[\![1\,;N]\!]$, considérons l'évènement :

E_i : « l'erreur n° i n'est pas corrigée à l'issue de n relectures ».

Cet évènement E_i peut lui-même s'écrire comme l'intersection des n évènements indépendants :

A_k : « l'erreur n° i n'est pas corrigée lors de la k^e relecture »,

et comme chaque A_k est de probabilité $1-p$ par hypothèse, on obtient :

$$P(E_i)=\prod_{k=1}^{n}P(A_k)=(1-p)^n\,.$$

L'évènement : « Le livre est entièrement corrigé après n relectures » est égal à $\bigcap_{i=1}^{N}\overline{E_i}$. Les évènements E_i étant indépendants, il en est de même des $\overline{E_i}$, donc la probabilité cherchée vaut : $\prod_{i=1}^{N}P(\overline{E_i})=\left(1-(1-p)^n\right)^N$.

Exemple

L'employé d'un hôtel doit ouvrir une chambre pour un client. Il dispose d'un trousseau avec les n clefs de l'hôtel mais celles-ci ont été mélangées, et il doit donc les essayer une par une au hasard (mais sans jamais essayer 2 fois la même).

Pour $k\in[\![1\,;n]\!]$, quelle est la probabilité qu'il ouvre la porte au k^e essai exactement ?

Notons A_k l'évènement :

« la porte s'ouvre au k^e essai ».

La probabilité cherchée est :

$$p_k=P(A_k\cap\overline{A_{k-1}}\cap\ldots\cap\overline{A_1})\,.$$

D'après la formule des probabilités composées,

$$p_k=P(\overline{A_1})P(\overline{A_2}\mid\overline{A_1})\cdots P(\overline{A_{k-1}}\mid\overline{A_1}\cap\ldots\cap\overline{A_{k-2}})P(A_k\mid\overline{A_1}\cap\ldots\cap\overline{A_{k-1}})\,.$$

On suppose bien sûr toutes les serrures différentes, de sorte que chaque clef a une probabilité égale d'ouvrir la porte. Donc, puisque les clefs déjà essayées sont mises de côté :

$$P(\overline{A_1}) = \frac{n-1}{n} \quad , \quad P\left(\overline{A_2} \,\middle|\, \overline{A_1}\right) = \frac{n-2}{n-1},$$

$$\dots$$

$$P(\overline{A_{k-1}} \mid \overline{A_1} \cap \dots \cap \overline{A_{k-2}}) = \frac{n-(k-1)}{n-(k-2)},$$

$$\text{et} \quad P(A_k \mid \overline{A_1} \cap \dots \cap \overline{A_{k-1}}) = \frac{1}{n-(k-1)}.$$

Par conséquent, la probabilité cherchée est :

$$p_k = \frac{1}{n-(k-1)} \times \frac{n-(k-1)}{n-(k-2)} \times \dots \times \frac{n-2}{n-1} \times \frac{n-1}{n} = \frac{1}{n}$$

(par télescopage).
Nous vous laissons interpréter ce résultat, qui était assez prévisible.

Exemple

Une urne contient n boules rouges et n boules blanches.
On tire les boules deux par deux, sans remise, jusqu'à vider l'urne.

Quelle est la probabilité que l'on tire une boule de chaque couleur à chaque tirage ?

Pour tout $k \in [\![1\,;n]\!]$, notons A_k l'évènement :

« on obtient une boule de chaque couleur au k^e tirage ».

L'évènement considéré dans l'énoncé est donc $\bigcap_{k=1}^{n} A_k$ dont on calcule la probabilité à l'aide de la formule des probabilités composées :

$$P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1) \prod_{k=1}^{n-1} \left(A_{k+1} \,\middle|\, \bigcap_{i=1}^{k} A_i\right).$$

Or $P(A_1) = \dfrac{\binom{n}{1}\binom{n}{1}}{\binom{2n}{2}} = \dfrac{n^2}{\binom{2n}{2}}$, *et plus généralement :*

$$P\left(A_{k+1} \,\middle|\, \bigcap_{i=1}^{k} A_i\right) = \frac{(n-k)^2}{\binom{2n-2k}{2}}.$$

On en déduit :

$$P\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=0}^{n-1} \frac{(n-k)^2}{\binom{2n-2k}{2}} = \prod_{k=0}^{n-1} \frac{2(n-k)^2}{(2n-2k)(2n-2k-1)} = \frac{2^n\,(n!)^2}{(2n)!}.$$

SF15.9 Écrire l'évènement comme intersection infinie d'évènements

Il s'agit ici d'écrire un évènement A sous la forme : $A = \bigcap_{k=1}^{+\infty} A_k$.

- Si la suite d'évènements (A_n) est décroissante, c'est-à-dire $A_{n+1} \subset A_n$, on a :
$$P(A) = \lim_{n\to+\infty} P(A_n).$$
- Sinon, on a : $P(A) = \lim_{n\to+\infty} P\left(\bigcap_{k=1}^{n} A_k\right)$.
- On peut aussi essayer de passer à l'évènement contraire, ce qui nous ramène au calcul de la probabilité d'une réunion infinie d'évènements.

Exemple

On lance un dé non pipé une infinité de fois.

Quelle est la probabilité que le 1 n'apparaisse jamais ?

Pour $n \geqslant 1$, notons A_n l'évènement :

« le 1 n'est pas apparu lors des n premiers tirages »,

et notons A l'évènement :

« le 1 n'apparaît jamais ».

On a alors $A = \bigcap_{n=1}^{+\infty} A_n$. D'autre part, la suite (A_n) est décroissante ($A_{n+1} \subset A_n$) puisque A_{n+1} implique A_n : si le 1 n'est pas apparu lors des $n+1$ premiers tirages, a fortiori il n'est pas apparu lors des n premiers !

On a donc $P(A) = \lim_{n\to+\infty} P(A_n)$. Mais $P(A_n) = \frac{5^n}{6^n}$: en effet, pour les n premiers lancers, l'univers est $\Omega = [\![1\,;6]\!]^n$ et l'évènement A_n est $[\![2\,;6]\!]^n$.

On en déduit $P(A) = 0$: l'évènement A est quasi-impossible.

> ✓ *Cependant l'évènement A n'est pas impossible. En effet, l'univers Ω qui décrit ici l'expérience est $[\![1\,;6]\!]^{\mathbb{N}}$, ensemble des suites à valeurs dans $[\![1\,;6]\!]$, et l'évènement A est $[\![2\,;6]\!]^{\mathbb{N}}$. A est donc loin d'être $\emptyset$ (il n'est même pas dénombrable !) mais cependant sa probabilité est nulle.*
> *Cet exemple illustre bien la complexité de l'étude des probabilités sur un univers non dénombrable.*

Exemple

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements mutuellement indépendants.

Montrer que la probabilité qu'aucun des A_n ne soit réalisé est inférieure à :
$$\exp\left(-\sum_{n=0}^{+\infty} P(A_n)\right).$$

La probabilité qu'aucun des A_n ne soit réalisé est :
$$P\left(\bigcap_{n=0}^{+\infty} \overline{A_n}\right) = \lim_{N\to+\infty} P\left(\bigcap_{n=0}^{N} \overline{A_n}\right).$$

Les évènements étant indépendants,

$$P\left(\bigcap_{n=0}^{N}\overline{A_n}\right)=\prod_{n=0}^{N}(1-P(A_n))\,.$$

Or $1-x\leqslant \mathrm{e}^{-x}$ *pour tout* $x\in\mathbb{R}$ *donc :*

$$P\left(\bigcap_{n=0}^{N}\overline{A_n}\right)\leqslant\prod_{n=0}^{N}\mathrm{e}^{-P(A_n)}=\exp\left(-\sum_{n=0}^{N}P(A_n)\right).$$

Quand $N\to+\infty$ *on obtient donc :*

$$P\left(\bigcap_{n=0}^{+\infty}\overline{A_n}\right)\leqslant\exp\left(-\sum_{n=0}^{+\infty}P(A_n)\right)$$

(où il faut bien sûr comprendre que, si la série $\sum P(A_n)$ *diverge (vers* $+\infty$*), l'exponentielle est nulle).*

SF15.10 Détecter 2 niveaux de hasard et appliquer la formule des probabilités totales

Nous considérons ici des situations où la réalisation d'un évènement (2e niveau de hasard) dépend d'évènements qui se sont passés auparavant (1er niveau).
Cette situation peut se produire lorsqu'on étudie des phénomènes qui évoluent au cours du temps et pour lesquels on connaît le lien entre deux réalisations successives. On applique alors la formule des probabilités totales avec un système complet d'évènements lié à l'instant $n-1$ à un évènement lié à l'instant n. On obtient alors des suites de probabilités définies par récurrence.

Mais on peut utiliser la formule des probabilités totales dans des situations plus variées ; il faut donc apprendre à reconnaître ces situations où interviennent deux niveaux de hasard, et y repérer les systèmes complets (ou quasi-complets) d'évènements.

Exemple

On étudie au cours du temps le fonctionnement d'un appareil obéissant aux règles suivantes :

- si l'appareil fonctionne à l'instant $n-1$ ($n\in\mathbb{N}^*$), il a la probabilité $\dfrac{1}{6}$ d'être en panne à l'instant n ;
- si l'appareil est en panne à l'instant $n-1$, il a la probabilité $\dfrac{2}{3}$ d'être en panne à l'instant n.

On note p_n la probabilité que l'appareil soit en état de marche à l'instant n.

Établir pour tout $n\in\mathbb{N}^*$ une relation entre p_n et p_{n-1}, et en déduire l'expression de p_n en fonction de p_0.

- *On note* F_n *l'évènement : « l'appareil fonctionne à l'instant* n *».*
  On applique alors la formule des probabilités totales avec le système complet d'évènements $(F_{n-1},\overline{F_{n-1}})$*. On a donc :*

$$P(F_n)=P(F_n\mid F_{n-1})P(F_{n-1})+P(F_n\mid\overline{F_{n-1}})P(\overline{F_{n-1}})$$

d'où $p_n=\dfrac{5}{6}p_{n-1}+\dfrac{1}{3}(1-p_{n-1})=\dfrac{1}{2}p_{n-1}+\dfrac{1}{3}$.

- *(p_n) est une suite arithmético-géométrique (voir* **[S8.5]***).*
  On résout alors l'équation $x = \frac{1}{2}x + \frac{1}{3}$ *:* $x = \frac{2}{3}$.
  On pose alors $q_n = p_n - \frac{2}{3}$. *On sait que* (q_n) *est une suite géométrique de raison* $\frac{1}{2}$ *donc* $q_n = \frac{1}{2^n} q_0 = \frac{1}{2^n}\left(p_0 - \frac{2}{3}\right)$.
  On en déduit : $p_n = \frac{1}{2^n}\left(p_0 - \frac{2}{3}\right) + \frac{2}{3}$.

Exemple

Une boite A contient deux jetons portant le numéro 0 et une boite B contient deux jetons portant le numéro 1.
On tire au hasard un jeton dans chaque boite et on les échange. On recommence cette opération n fois. On s'intéresse à la somme des numéros des jetons contenus dans la boîte A à l'instant $t = n$.
Pour cela, on introduit les évènements :

– A_n : « la somme des n^{os} des jetons de la boîte A à l'instant $t = n$ vaut 0 » ;
– B_n : « la somme des n^{os} des jetons de la boîte A à l'instant $t = n$ vaut 1 » ;
– C_n : « la somme des n^{os} des jetons de la boîte A à l'instant $t = n$ vaut 2 ».

On pose également $a_n = P(A_n)$, $b_n = P(B_n)$ et $c_n = P(C_n)$.

1) Calculer a_0, b_0, c_0, a_1, b_1, c_1.

2) Exprimer a_{n+1}, b_{n+1} et c_{n+1} en fonction de a_n, b_n, c_n.

3) En déduire a_n, b_n et c_n en fonction de n, puis leurs limites quand $n \to +\infty$.

1) *De façon claire :* $a_0 = 1$, $b_0 = 0$, $c_0 = 0$ *et* $a_1 = 0$, $b_1 = 1$, $c_1 = 0$.

2) *Le système* $\{A_n, B_n, C_n\}$ *est un système complet d'évènements. Par la formule des probabilités totales :*

$$\begin{aligned} a_{n+1} &= P(A_{n+1}) = P(A_{n+1} \cap A_n) + P(A_{n+1} \cap B_n) + P(A_{n+1} \cap C_n) \\ &= \underbrace{P\left(A_{n+1}|A_n\right)}_{=0} P(A_n) + P\left(A_{n+1}|B_n\right) P(B_n) + \underbrace{P\left(A_{n+1}|C_n\right)}_{=0} P(C_n) \\ &= P\left(A_{n+1} \mid B_n\right) P(B_n) = \frac{1}{4}\, b_n \,. \end{aligned}$$

$$\begin{aligned} b_{n+1} &= P(B_{n+1}) \\ &= \underbrace{P\left(B_{n+1}|A_n\right)}_{=1} P(A_n) + P\left(B_{n+1}|B_n\right) P(B_n) + \underbrace{P\left(B_{n+1}|C_n\right)}_{=1} P(C_n) \\ &= P(A_n) + P\left(B_{n+1}|B_n\right) P(B_n) + P(C_n) = a_n + \frac{1}{2}\, b_n + c_n \,. \end{aligned}$$

$$\begin{aligned} c_{n+1} &= P(C_{n+1}) \\ &= \underbrace{P\left(C_{n+1}|A_n\right)}_{=0} P(A_n) + P\left(C_{n+1}|B_n\right) P(B_n) + \underbrace{P\left(C_{n+1}|C_n\right)}_{=0} P(C_n) \\ &= P\left(C_{n+1} \mid B_n\right) P(B_n) = \frac{1}{4}\, b_n \,. \end{aligned}$$

Toutes ces probabilités ont été calculées en utilisant le fait que si l'on connaît la somme des numéros des jetons contenus dans la boîte A, on connaît en fait le contenu des deux boîtes.

3) *On a ainsi un système linéaire de suites récurrentes, comme nous en avons déjà étudié en* **[SF4.20]**. *On pourrait donc calculer a_n, b_n et c_n en calculant les puissances de la matrice* $A = \begin{pmatrix} 0 & 1/4 & 0 \\ 1 & 1/2 & 1 \\ 0 & 1/4 & 0 \end{pmatrix}$.

✓ *Une telle matrice est dite* stochastique. *La somme des valeurs dans chaque colonne est égale à* 1. *Sauriez-vous expliquer pourquoi ?*

Mais il y a plus rapide. En effet, à partir des relations précédentes, on obtient pour tout entier n :

$$b_{n+2} = a_{n+1} + \frac{1}{2}b_{n+1} + c_{n+1} = \frac{1}{4}b_n + \frac{1}{2}b_{n+1} + \frac{1}{4}b_n = \frac{1}{2}b_{n+1} + \frac{1}{2}b_n\,.$$

(b_n) est donc une suite récurrente linéaire d'ordre 2. L'équation caractéristique de la récurrence est : $x^2 - \frac{1}{2}x - \frac{1}{2} = 0$, de racines 1 et $-\frac{1}{2}$. On en déduit, compte tenu des valeurs initiales : $\forall\, n \in \mathbb{N}$, $b_n = -\frac{2}{3}\left(-\frac{1}{2}\right)^n + \frac{2}{3}$,

d'où l'on tire, pour $n \geqslant 1$: $a_n = c_n = \frac{1}{4}b_{n-1} = \frac{1}{4}\left(-\frac{2}{3}\left(-\frac{1}{2}\right)^{n-1} + \frac{2}{3}\right)$.

On en déduit : $\lim\limits_{n\to+\infty} a_n = \frac{1}{6}$, $\lim\limits_{n\to+\infty} b_n = \frac{2}{3}$ *et* $\lim\limits_{n\to+\infty} c_n = \frac{1}{6}$.

Exemple

On dispose de $N+1$ urnes numérotées de 0 à N. L'urne numéro k contient k boules blanches et $N-k$ boules noires.
On choisit une urne au hasard, et on y tire des boules au hasard, avec remise.

1) On tire n boules. Déterminer la probabilité que ces n boules soient blanches. Quelle est la limite de cette probabilité lorsque N tend vers $+\infty$?

2) On tire $n+1$ boules. Sachant que les n premières ont été blanches, quelle est la probabilité que la $(n+1)^{\text{e}}$ le soit aussi ? Quelle est la limite de cette probabilité lorsque N tend vers $+\infty$?

1) *Le 1er niveau de hasard est ici le choix de l'urne. Si l'on considère les évènements :*

$$U_k \text{ : « le tirage s'est fait dans l'urne n}^o\, k \text{ »},$$

le système $\{U_0, \ldots, U_N\}$ forme un système complet d'évènements.
Notons $A_{n,N}$ l'évènement : « les n boules tirées sont blanches ». D'après la formule des probabilités totales :

$$P(A_{n,N}) = \sum_{k=0}^{N} P\left(A_{n,N} \mid U_k\right) P(U_k) = \sum_{k=0}^{N} \frac{1}{N+1}\left(\frac{k}{N}\right)^n.$$

On reconnaît ici (presque) une somme de Riemann associée à la fonction continue $x \mapsto x^n$ sur l'intervalle $[0\,;1]$ découpé en N intervalles réguliers.

Plus précisément, $P(A_{n,N}) = \dfrac{N}{N+1} \cdot \dfrac{1}{N} \displaystyle\sum_{k=0}^{N} \left(\frac{k}{N}\right)^n$, *et l'on sait que :*

$$\lim_{N\to+\infty} \left(\frac{1}{N} \sum_{k=0}^{N} \left(\frac{k}{N}\right)^n \right) = \int_0^1 x^n \,\mathrm{d}x = \frac{1}{n+1} \cdot$$

On a donc $\displaystyle\lim_{N\to+\infty} P(A_{n,N}) = \frac{1}{n+1} \cdot$

2) *Si l'on note* B *l'évènement : « la* $(n+1)^e$ *boule tirée est blanche », on veut calculer* $P(B \mid A_{n,N})$. *On a donc :*

$$P(B \mid A_{n,N}) = \frac{P(A_{n,N} \cap B)}{P(A_{n,N})} = \frac{P(A_{n+1,N})}{P(A_{n,N})},$$

puisque, clairement, $A_{n,N} \cap B = A_{n+1,N}$. *Ainsi :*

$$P(B \mid A_{n,N}) = \frac{\displaystyle\sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n+1}}{\displaystyle\sum_{k=0}^{N} \left(\frac{k}{N}\right)^{n}} = \frac{1^{n+1} + 2^{n+1} + \cdots + N^{n+1}}{N(1^n + 2^n + \cdots + N^n)},$$

et, compte tenu des calculs précédents : $\displaystyle\lim_{N\to+\infty} P(B \mid A_{n,N}) = \frac{n+1}{n+2} \cdot$

Ce dernier résultat porte le nom de loi de succession de Laplace.

■ Comment calculer une probabilité conditionnelle ?

Le calcul de probabilités conditionnelles intervient naturellement lorsqu'il s'agit de répondre à des questions directes comme :

« sachant qu'un évènement A s'est produit,
calculer la probabilité de l'évènement B »,

ainsi que nous allons le voir dans les exemples suivants.
Mais il ne faut pas oublier qu'elles interviennent très fréquemment dans de nombreux autres problèmes, dès lors qu'on est amené à utiliser la formule des probabilités totales ou celle des probabilités composées, comme nous l'avons vu dans plusieurs exemples auparavant.

SF15.11 Utiliser la définition

On utilise ici simplement la définition : $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} \cdot$

Exemple

Une commode avec n tiroirs contient une paire de chaussettes avec la probabilité p (chaque tiroir ayant la même chance de la contenir).
On ouvre $n-1$ tiroirs et ils sont vides.
Quelle est la probabilité que le dernier tiroir contienne la paire de chaussettes ?

Pour $k \in [\![1\,;n]\!]$, *notons* A_k *l'évènement :*

A_k : « la paire de chaussette est dans le k^e tiroir ».

La probabilité cherchée est donc

$$P\left(A_n \mid \overline{A_1} \cap \ldots \cap \overline{A_{n-1}}\right) = \frac{P(A_n \cap \overline{A_1} \cap \ldots \cap \overline{A_{n-1}})}{P(\overline{A_1} \cap \ldots \cap \overline{A_{n-1}})}.$$

Mais $A_n \cap \overline{A_1} \cap \ldots \cap \overline{A_{n-1}} = A_n$ car si la paire se trouve dans le dernier tiroir, cela implique qu'elle n'était pas dans les précédents. De plus,

$$P(\overline{A_1} \cap \ldots \cap \overline{A_{n-1}}) = 1 - P(A_1 \cup \ldots \cup A_{n-1})$$

et comme les évènements A_k sont disjoints (il y a au plus une paire de chaussettes) et de même probabilité $\frac{p}{n}$, on obtient :

$$P(\overline{A_1} \cap \ldots \cap \overline{A_{n-1}}) = 1 - \sum_{k=1}^{n-1} P(A_k) = 1 - (n-1)\frac{p}{n}.$$

La probabilité cherchée est donc finalement égale à :

$$\frac{P(A_n)}{P(\overline{A_1} \cap \ldots \cap \overline{A_{n-1}})} = \frac{p}{n-(n-1)p}.$$

SF15.12 Utiliser la formule de Bayes

Il ne faut pas oublier que la formule donnant une probabilité conditionnelle permet aussi de calculer $P(A \mid B)$ si l'on connaît $P(B \mid A)$:

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}.$$

Sous cette forme, elle permet de déterminer la probabilité de la « cause » A lorsque la « conséquence » B s'est réalisée, si bien sûr on connaît les probabilités de tous les autres évènements qui sont intervenus.
Lorsque l'on calcule le dénominateur $P(B)$ par la formule des probabilités totales, lorsqu'il y a plusieurs causes, on obtient la formule de Bayes rappelée en **[S15.4]**.

Exemple

Une urne contient une boule blanche et une boule noire.
On effectue un premier tirage ; si la boule obtenue est blanche on la remet dans l'urne et on ajoute une boule blanche ; si elle est noire, on la remet simplement dans l'urne.
Lors d'un second tirage, on tire une boule blanche. Quelle est la probabilité que la première boule tirée ait été blanche ?

Pour $i \in \{1,2\}$, on note B_i (respectivement N_i) l'évènement :
« la i^e boule est blanche (respectivement noire) ».

La probabilité cherchée est $P(B1 \mid B_2)$ qui est égale à :

$$P(B_1 \mid B_2) = \frac{P(B_2 \mid B_1)\,P(B_1)}{P(B_2)}.$$

On connaît $P(B_1) = \frac{1}{2}$ et $P(B_2 \mid B_1) = \frac{2}{3}$. On calcule ensuite $P(B_2)$ par la formule des probabilités totales avec le système complet d'évènements

$\{B_1,N_1\}$:

$$P(B_2) = P(B_2 \mid B_1)\,P(B_1) + P(B_2 \mid N_1)\,P(N_1) = \frac{2}{3}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{7}{12}\cdot$$

En conclusion : $P(B_1 \mid B_2) = \dfrac{\frac{2}{3}\cdot\frac{1}{2}}{\frac{7}{12}} = \frac{4}{7}\cdot$

✓ Vous remarquerez que l'on n'a pas employé directement la formule de Bayes : on l'a retrouvée en utilisant les autres formules.

■ Comment prouver que des évènements sont indépendants ?

SF15.13 Revenir à la définition

Lors de l'étude de l'indépendance d'évènements, il faut absolument éviter toute démarche « intuitive » ; en effet, la notion d'indépendance est intimement liée à la probabilité qui a été définie, et elle doit donc être vérifiée par le calcul.
Les deux exemples qui suivent illustrent la difficulté de cette notion, et son caractère souvent contraire à l'intuition immédiate.

Exemple

On considère une famille ayant n enfants ($n \geqslant 2$), et les évènements :

A : « il y a des garçons et des filles »

B : « il y a des garçons et au plus une fille ».

En supposant que la probabilité d'avoir une fille et celle d'avoir un garçon sont égales, déterminer la probabilité des évènements A et B. Ces évènements sont-ils indépendants ?

L'univers Ω est ici l'ensemble $\{F,G\}^n$, muni de la probabilité uniforme.

- *L'évènement contraire de A est l'évènement :*

  $\overline{A}$ *: « il n'y a que des garçons ou que des filles »,*

  lui-même étant réunion disjointe des deux évènements $\{F\}^n$ et $\{G\}^n$. Puisque $\mathrm{Card}(\Omega) = 2^n$ *et* $\mathrm{Card}(\overline{A}) = 2$*, on a donc :*

  $$P(A) = 1 - \frac{2}{2^n} = 1 - \frac{1}{2^{n-1}}\cdot$$

- *L'évènement B est la réunion des deux évènements disjoints :*

  B_1 *: « il n'y a que des garçons »*

  et B_2 *: « il y a $(n-1)$ garçons et 1 fille ».*

  $\mathrm{Card}(B_1) = 1$ *et* $card(B_2) = n$ *puisque B_2 est l'ensemble des n-uplets $(G,\ldots,G,F,G,\ldots,G)$. Donc :* $P(B) = \dfrac{n+1}{2^n}\cdot$

- *Avec les notations ci-dessus, on a* $A \cap B = B_2$*, donc* $P(A \cap B) = \dfrac{n}{2^n}\cdot$

  Les évènements A et B sont indépendants si et seulement si :

$$P(A\cap B)=P(A)P(B)\,,\quad \textit{soit}\quad \frac{n}{2^n}=\left(1-\frac{1}{2^{n-1}}\right)\left(\frac{n+1}{2^n}\right).$$

Cela équivaut à $n+1=2^{n-1}$, ce qui n'est possible que si $n=3$.

Exemple

On suppose qu'il existe une probabilité P sur l'espace probabilisable $(\mathbb{N},\mathcal{P}(\mathbb{N}))$ telle que pour tout entier $k\in\mathbb{N}^*$, la probabilité pour un entier d'être multiple de k est égale à $\frac{1}{k}$.

Pour tout $k\in\mathbb{N}^*$, on notera $A_k=k\mathbb{N}$ l'ensemble des multiples de k, et $\mathbb{P}$ l'ensemble des nombres premiers.

1) Montrer que les évènements A_p pour $p\in\mathbb{P}$ sont indépendants.

2) En déduire une contradiction.

L'hypothèse de l'énoncé s'écrit : $P(A_k)=\frac{1}{k}$.

1) *Si $p_1,\dots,p_k$ sont des nombres premiers distincts, alors $A_{p_1}\cap\dots\cap A_{p_k}$ est l'ensemble des entiers multiples de tous les p_i donc c'est exactement l'ensemble des entiers multiples du produit $N=p_1p_2\dots p_k$.*

Donc $P(A_{p_1}\cap\dots\cap A_{p_k})=P(A_N)=\frac{1}{N}=\prod_{i=1}^{k}\frac{1}{p_i}=\prod_{i=1}^{k}P(A_{p_i})\,,$

et, par définition, la famille $(A_p)_{p\in\mathbb{P}}$ est formée d'évènements indépendants.

2) *Supposons désormais les nombres premiers p_i rangés en ordre croissant ($p_1=2$, $p_2=3$, $p_3=5$, …).*

Les A_{p_i} étant indépendants, il en est de même des évènements $\overline{A_{p_i}}$ donc pour tout $n\in\mathbb{N}^*$:

$$P\left(\bigcap_{i=1}^{n}\overline{A_{p_i}}\right)=\prod_{i=1}^{n}P(\overline{A_{p_i}})=\prod_{i=1}^{n}\left(1-\frac{1}{p_i}\right).$$

On en déduit : $P\left(\bigcap_{i=1}^{+\infty}\overline{A_{p_i}}\right)=\lim_{n\to+\infty}\prod_{i=1}^{n}\left(1-\frac{1}{p_i}\right).$

Nous admettrons que la série $\sum_{i\geqslant 1}\frac{1}{p_i}$ diverge (il existe de nombreuses démonstrations de ce résultat) ; puisque $\ln\left(1-\frac{1}{p_i}\right)\underset{i\to+\infty}{\sim}-\frac{1}{p_i}$, on en déduit que $\lim_{n\to+\infty}\sum_{i=1}^{n}\ln\left(1-\frac{1}{p_i}\right)=-\infty$ donc que $\lim_{n\to+\infty}\prod_{i=1}^{n}\left(1-\frac{1}{p_i}\right)=0$.

Ainsi : $P\left(\bigcap_{i=1}^{+\infty}\overline{A_{p_i}}\right)=0.$

On démontre exactement de la même façon que, pour tout entier $k\geqslant 1$, $P\left(\bigcap_{i=k}^{+\infty}\overline{A_{p_i}}\right)=0$ *(il suffit de changer « $i=1$ » en « $i=k$ » dans les calculs précédents).*

Or puisque $p_i \geqslant i+1$ pour tout i, l'ensemble d'entiers $[\![1\,;k]\!]$ est inclus dans $\bigcap_{i=k}^{+\infty} \overline{A_{p_i}}$ (car $1 \leqslant x \leqslant k \implies 1 \leqslant x < p_k$ donc x ne peut être multiple d'aucun p_i avec $i \geqslant k$).

On aurait donc, pour tout entier $k \geqslant 1$, $P([\![1\,;k]\!]) = 0$ d'où par continuité croissante $P(\mathbb{N}^*) = 0$ ce qui est absurde et achève la preuve.

✓ Ainsi, il n'existe pas de « vraie » probabilité sur $\mathbb{N}$ qui corresponde à l'idée intuitive que l'on peut en avoir !

SF15.14 Se ramener à des évènements indépendants connus

Dans beaucoup de cas, l'indépendance d'évènements est directement supposée par l'énoncé, ou découle des conditions de l'expérience aléatoire.

On peut alors utiliser la propriété suivante (voir **[S15.5]**) : si $(A_i)_{i\in I}$ est une famille d'évènements indépendants, toute famille formée à partir des A_i ou de leurs complémentaires est aussi formée d'évènements indépendants.

Nous avons déjà eu l'occasion d'utiliser cette propriété dans des exemples précédents, dont celui ci-dessus.

À vous de jouer !

Exercice 15.1 **SF15.5**

Soient $(\Omega,\mathcal{A},P)$ un espace probabilisé, et A, B deux évènements dans $\mathcal{A}$.

Montrer que :

$$|P(A\cap B) - P(A)P(B)| \leqslant \frac{1}{4}\cdot$$

Exercice 15.2 **SF15.3** **SF15.4**

On dispose de N tickets de loterie, dont n sont gagnants.

Quelle est la probabilité qu'une personne achetant n tickets en obtienne au moins un gagnant ?

Exercice 15.3 **SF15.8** **SF15.13**

Pour tout entier $n \in \mathbb{N}^*$, on note $\varphi(n)$ le nombre d'entiers entre 1 et n qui sont premiers avec n.
Il s'agit de l'indicateur d'Euler $\varphi(n)$ (cf. **[S1.6]**).
On munit l'ensemble $\Omega = [\![1\,;n]\!]$ de la loi de probabilité uniforme.

1) Si d est un diviseur de n, on note D_d l'ensemble des multiples de d dans Ω. Calculer $P(D_d)$.

2) On note p_1,p_2,...,p_r les facteurs premiers de n.

a) Montrer que les évènements D_{p_1}, D_{p_2}, ..., D_{p_r} sont mutuellement indépendants.

b) Retrouver ainsi la formule :

$$\varphi(n) = n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_r}\right).$$

Exercice 15.4

SF15.7 **SF15.9**

Soit $(\Omega,\mathcal{A},P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une famille quelconque d'évènements.

1) On suppose ici que la série $\sum_{n\in\mathbb{N}} P(A_n)$ est convergente.

En considérant l'évènement :

$$B = \bigcap_{n=0}^{+\infty}\left(\bigcup_{k=n}^{+\infty} A_k\right)$$

et son complémentaire, montrer qu'il est quasi certain que seul un nombre fini d'évènements A_k se réalise simultanément.

2) On suppose maintenant les évènements A_n mutuellement indépendants, et la série $\sum_{n\in\mathbb{N}} P(A_n)$ divergente.

Montrer qu'il est quasi certain qu'une infinité d'évènements A_k se réalisent simultanément.

Exercice 15.5

SF15.8 **SF15.11** **SF15.12**

Au village de St-Amand, il fait beau 5 jours par semaine en moyenne. Pour connaître le temps qu'il fera, Mr Lambda a deux sources d'informations :

– la météo nationale à la radio, qui a raison à 95% ;

– le vieux paysan du village, qui a raison à 90%.

La météo nationale annonce du mauvais temps, mais le vieux paysan prédit du soleil.

Mr Lambda doit-il prendre son parapluie ?

Exercice 15.6

SF15.10

Deux urnes U_1 et U_2 contiennent des boules blanches et noires.
L'urne U_1 contient une proportion p_1 de boules blanches et l'urne U_2 en contient une proportion p_2.
On effectue des tirages avec remise dans ces urnes de la façon suivante.

– On effectue un premier tirage dans une urne choisie au hasard et on remet la boule obtenue dans son urne d'origine.
Si l'on obtient une boule blanche (respectivement noire), le 2e tirage se fait dans U_1 (respectivement U_2).

– Au i^{e} tirage, si la boule obtenue est blanche (respectivement noire), le $(i+1)^{\text{e}}$ tirage se fait dans U_1 (respectivement U_2).

On note B_i l'évènement : « on obtient une boule blanche au i^{e} tirage ».

1) Calculer $P(B_1)$.

2) Exprimer $P(B_n)$ en fonction de $P(B_{n-1})$.

3) Montrer que la suite $\big(P(B_n)\big)$ converge et déterminer sa limite.

Exercice 15.7 **SF15.7** **SF15.10** **SF15.11**

On admet que dans une population la probabilité p_n pour une famille d'avoir exactement n enfants est donnée par :

$$\forall\, n \geqslant 1,\ p_n = \alpha p^n$$

où $p \in]0\,;1[$ et $p(1+\alpha) < 1$.
On suppose de plus que les naissances d'une fille ou d'un garçon sont équiprobables.

1) Quelle est la probabilité qu'une famille n'ait pas d'enfants ?

2) Pour $k \geqslant 1$, calculer la probabilité qu'une famille ait exactement k garçons.

3) Étant donnée une famille ayant au moins un garçon, quelle est la probabilité qu'elle en ait au moins deux ?

Exercice 15.8 **SF15.7** **SF15.8** **SF15.11**

Une urne contient initialement b boules blanches et r boules rouges. On effectue des tirages dans cette urne de la façon suivante : si à un rang quelconque on obtient une boule rouge, on la remet dans l'urne, et si l'on obtient une boule blanche, on ne la remet pas.

1) Quelle est la probabilité de tirer exactement une boule blanche au cours des n premiers tirages ?

2) Sachant qu'au cours des n premiers tirages on a tiré exactement une boule blanche, quelle est la probabilité qu'elle ait été tirée en dernier ?

Thème 16 - Variables aléatoires discrètes

■ Comment déterminer la loi d'une variable aléatoire ?

SF16.1 Reconnaître directement une loi classique

Pour reconnaître une des lois qui figurent au programme, il est indispensable de bien connaître les situations-type d'application qui ont été rappelées en **[S16.2]**. La seule difficulté est alors que, dans certains problèmes, la façon dont est formulé l'énoncé masque la présence d'une loi classique.

Exemple

n personnes lancent chacune une pièce de monnaie équilibrée ($n > 2$).
Une personne est dite gagnante si elle obtient le contraire de toutes les autres.
On appelle X le nombre de parties nécessaires pour qu'il y ait un gagnant.
Déterminer la loi de X, son espérance, sa variance.

Notons pour abréger P pour « pile » et F pour « face ».
L'univers décrivant l'expérience est ici $\Omega = \{P,F\}^n$, de cardinal 2^n.
Une situation gagnante est un n-uplet formé d'un seul P et de $(n-1)$ F (n possibilités) ou un n-uplet formé d'un seul F et de $(n-1)$ P (n possibilités).
Pour chaque partie, la probabilité qu'il y ait un gagnant est donc :

$$p = \frac{2n}{2^n} = \frac{n}{2^{n-1}}\cdot$$

La variable X correspond donc au nombre d'épreuves nécessaires pour obtenir un succès, la probabilité de succès à chaque épreuve étant égale à p ; ainsi, X suit la loi géométrique de paramètre p.
On a alors, directement d'après le cours :

$$E(X) = \frac{1}{p} \quad \text{et} \quad V(X) = \frac{1-p}{p^2}\cdot$$

Exemple

Le nombre X d'œufs pondus par un insecte suit la loi de Poisson de paramètre λ. La probabilité que chaque œuf se développe est p.

On note Y le nombre de survivants. Quelle est la loi de Y ?

Le fait qu'un œuf survive (=succès) est une expérience de Bernoulli de paramètre p. Lorsqu'il y a n œufs, le nombre de survivants (=nombre de succès) suit donc la loi binomiale de paramètres n et p.

Ainsi, pour $0 \leqslant k \leqslant n$, $P(Y = k \mid X = n) = \binom{n}{k} p^k (1-p)^{n-k}$.

Y est à valeurs dans $\mathbb{N}$. Par la formule des probabilités totales on a, pour tout $k \in \mathbb{N}$:

$$P(Y=k)=\sum_{n=k}^{+\infty} P\left(Y=k|X=n\right)P(X=n)=\sum_{n=k}^{+\infty}\binom{n}{k}p^k(1-p)^{n-k}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$$
$$=\mathrm{e}^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{n=k}^{+\infty}\frac{(\lambda(1-p))^{n-k}}{(n-k)!}=\mathrm{e}^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{i=0}^{+\infty}\frac{(\lambda(1-p))^i}{i!}$$
$$=\mathrm{e}^{-\lambda}\frac{(\lambda p)^k}{k!}\mathrm{e}^{\lambda(1-p)}=\mathrm{e}^{-\lambda p}\frac{(\lambda p)^k}{k!}\,.$$

Y suit donc la loi de Poisson de paramètre λp.

SF16.2 Déterminer $X(\Omega)$ et calculer $P(X=x)$ pour tout $x\in X(\Omega)$

Il s'agit ici de revenir à la définition de la loi d'une variable aléatoire X. On est ainsi ramené à calculer la probabilité des évènements $P(X=x)$, ce qui peut se faire par les méthodes exposées dans le chapitre précédent.

Dans certains cas, il est plus facile de calculer la probabilité des évènements $(X\leqslant x)$ ou $(X<x)$ (ou leurs contraires), puis d'en déduire celle de l'évènement $(X=x)$ (cala revient à considérer la fonction de répartition de X).

✓ Dans tous les cas, on n'oubliera pas de déterminer en premier l'ensemble $X(\Omega)$ des valeurs prises par X.

Exemple

N urnes contiennent chacune des boules numérotées de 1 à n. On tire au hasard une boule dans chaque urne, et on appelle X le plus grand des numéros obtenus.

Trouver la loi de X.

L'ensemble des valeurs prises par X est $[\![1\,;n]\!]$.
Pour chercher la loi de X, il est bien plus commode de déterminer d'abord les probabilités $P(X\leqslant k)$ pour $k\in[\![1\,;n]\!]$.

Soit donc $k\in[\![1\,;n]\!]$. Pour $i\in[\![1\,;n]\!]$ on considère l'évènement :

$$A_i\ :\ \text{« on tire un numéro} \leqslant k \text{ dans l'urne n}^o\, i\ \text{»}.$$

Les différents tirages étant indépendants, on a :

$$P(X\leqslant k)=P(A_1\cap\ldots\cap A_N)=\prod_{i=1}^{N}P(A_i)=\left(\frac{k}{n}\right)^N.$$

On a ensuite directement :

$$\forall\, k\in[\![2\,;n]\!],\ P(X=k)=P(X\leqslant k)-P(X\leqslant k-1))=\frac{k^N-(k-1)^N}{n^N}\,,$$

et cette formule reste valable pour $k=1$.

Exemple

On effectue une série de lancers simultanés de deux pièces, chacune donnant « pile » avec la même probabilité p. On appelle X la variable aléatoire correspondant au nombre de lancers nécessaires pour obtenir au moins un « pile » sur l'une des pièces.

Déterminer la loi de X.

Si l'on note X_1 (respectivement X_2) le nombre de lancers nécessaires pour obtenir « pile » avec la 1re (respectivement 2e) pièce, alors $X = \min(X_1, X_2)$. X_1 et X_2 suivant des lois géométriques de paramètre p, X est à valeurs dans $\mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$:

$$P(X \geqslant n) = P\left(X_1 \geqslant n \cap X_2 \geqslant n\right) = P(X_1 \geqslant n)P(X_2 \geqslant n)\,.$$

$$\text{Or :}\quad P(X_1 \geqslant n) = \sum_{k=n}^{+\infty} P(X_1 = k) = \sum_{k=n}^{+\infty} (1-p)^{k-1}p$$
$$= p(1-p)^{n-1} \cdot \frac{1}{1-(1-p)} = (1-p)^{n-1}\,,$$

et puisque $P(X_1 \geqslant n) = P(X_2 \geqslant n)$ on obtient :

$$P(X \geqslant n) = (1-p)^{2n-2}\,.$$

✓ *On pouvait aussi obtenir directement la valeur de $P(X_1 \geqslant n)$ par un raisonnement probabiliste ; en effet, la probabilité $P(X_1 \geqslant n)$ est égale à $P(X_1 > n-1)$, c'est-à-dire à la probabilité qu'il n'y ait eu que des échecs lors des $n-1$ premiers lancers ; elle vaut donc $(1-p)^{n-1}$.*

Finalement,

$$P(X = n) = P(X \geqslant n) - P(X \geqslant n+1)$$
$$= (1-p)^{2n-2} - (1-p)^{2n} = \left((1-p)^2\right)^{n-1}(2p-p^2)\,.$$

✓ *On peut remarquer que, puisque $1-(1-p)^2 = 2p-p^2$, X suit la loi géométrique de paramètre $2p-p^2$. Sauriez-vous expliquer pourquoi ce résultat était prévisible ?*

SF16.3 Exprimer X en fonction d'une autre variable aléatoire

Lorsque le résultat d'une expérience dépend d'une expérience précédente, la loi d'une variable aléatoire décrivant la seconde expérience peut être fonction d'une variable aléatoire liée à la première expérience ; si l'énoncé ne le fait pas, c'est à vous d'introduire cette variable aléatoire.

Exemple

Une urne contient des boules blanches et des boules noires, les boules blanches étant en proportion $p \in]0\,;1[$.
On effectue 2 tirages successifs avec remise, et on note X le rang d'apparition de la 2e boule blanche.

Déterminer la loi de X et son espérance.

• *Notons déjà qu'il faut au moins 2 tirages pour obtenir 2 boules blanches, donc $X(\Omega) = \mathbb{N} \setminus \{0,1\}$.*
Appelons Y la variable aléatoire égale au rang d'apparition de la 1re boule blanche. On sait que Y suit la loi géométrique de paramètre p. En utilisant la formule des probabilités totales avec le système complet d'évènements $\{(Y=i),\, i \in \mathbb{N}^*\}$ on a, pour tout entier $j \geqslant 2$:

$$P(X = j) = \sum_{i=1}^{+\infty} P\left(X = j \cap Y = i\right)\,.$$

Mais puisque la seconde boule blanche ne peut apparaître qu'après la première (!), l'évènement $(X=j)\cap(Y=i)$ est impossible si $i \geqslant j$. Donc :

$$P(X=j)=\sum_{i=1}^{j-1} P\left(X=j\cap Y=i\right)=\sum_{i=1}^{j-1} P\left(X=j \mid Y=i\right) P(Y=i)\,.$$

De plus, les deux tirages étant indépendants, on remet les compteurs à zéro après le i^e tirage, donc $P\left(X=j \mid Y=i\right)=P(Y=j-i)$.

Finalement on obtient, en utilisant la définition d'une loi géométrique :

$$\begin{aligned} P(X=j) &= \sum_{i=1}^{j-1}(1-p)^{j-i-1}p\times(1-p)^{i-1}p = p^2\sum_{i=1}^{j-1}(1-p)^{j-2} \\ &= (j-1)p^2(1-p)^{j-2}\,. \end{aligned}$$

- *Le calcul de l'espérance de X est classique.*

$$E(X)=\sum_{j=2}^{+\infty} jP(X=j)=p^2\sum_{j=2}^{+\infty} j(j-1)(1-p)^{j-2}\,.$$

Pour tout $x\in\,]-1\,;1[$ on a, d'après le cours sur les séries entières :

$$\sum_{n=2}^{+\infty} n(n-1)x^{n-2}=\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\sum_{n=0}^{+\infty} x^n\right)=\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{1}{1-x}\right)=\frac{2}{(1-x)^3}$$

et finalement :

$$E(X)=p^2\frac{2}{(1-(1-p))^3}=\frac{2}{p}\,.$$

✓ La méthode utilisée dans cet exemple ressemble beaucoup à celle présentée dans l'exemple suivant. Mais ici, c'était à nous d'introduire une variable aléatoire convenable. De plus, cet exemple montre que, finalement, les couples de variables aléatoires apparaissent de façon naturelle dans de nombreuses situations, sans que forcément l'énoncé ne les mentionne.

SF16.4 Utiliser la loi d'un couple

Pour déterminer les lois de deux variables X et Y, il est parfois plus facile de déterminer la loi du couple (X,Y), puis d'en déduire les lois marginales par les formules rappelées en **[S16.4]**.

Exemple

Une urne contient 2 boules blanches et $n-2$ boules rouges. On effectue n tirages sans remise dans cette urne. On appelle X le rang de sortie de la première boule blanche et Y le rang de sortie de la deuxième boule blanche.

1) Déterminer la loi du couple (X,Y).

2) En déduire la loi de Y.

1) *On a ici $X(\Omega)=[\![1\,;n-1]\!]$ et $Y(\Omega)=[\![2\,;n]\!]$.*

On notera R_k (respectivement B_k) l'évènement :

« on a obtenu une boule rouge (respectivement blanche) au k^e tirage ».

Soient $i \in X(\Omega)$ et $j \in Y(\Omega)$.

– *Si $i \geqslant j$ alors $P\big(X = i \cap Y = j\big) = 0$.*

– *Si $i < j$ alors on peut écrire*

$$(X = i) \cap (Y = j) = R_1 \cap \cdots \cap R_{i-1} \cap B_i \cap R_{i+1} \cap \cdots \cap R_{j-1} \cap B_j\,,$$

et donc d'après la formule des probabilités composées :

$$\begin{aligned} P\big((X = i) \cap (Y = j)\big) &= \frac{n-2}{n} \times \frac{n-3}{n-1} \times \cdots \times \frac{n-2-(i-2)}{n-(i-2)} \times \frac{2}{n-(i-1)} \\ &\quad \times \frac{n-2-(i-1)}{n-i} \times \cdots \times \frac{n-2-(j-3)}{n-(j-2)} \times \frac{1}{n-(j-1)} \\ &= \frac{(n-2)(n-3)\cdots(n-i) \times 2 \times (n-i-1)\cdots(n-j+1)}{n(n-1)(n-2)\cdots(n-j+1)} \\ &= \frac{2}{n(n-1)}\,. \end{aligned}$$

2) *D'après la formule des probabilités totales utilisée avec le système complet d'évènements $\{(X = i),\ 1 \leqslant i \leqslant n-1\}$, on a pour tout $j \in [\![2\,;n]\!]$:*

$$\begin{aligned} P(Y = j) &= \sum_{i=1}^{n-1} P\big((X = i) \cap (Y = j)\big) \\ &= \sum_{i=1}^{j-1} P\big((X = i) \cap (Y = j)\big) + \sum_{i=j}^{n-1} \underbrace{P\big((X = i) \cap (Y = j)\big)}_{=0} \\ &= \sum_{i=1}^{j-1} \frac{2}{n(n-1)} = \frac{2(j-1)}{n(n-1)}\,. \end{aligned}$$

✓ Après tous ces calculs, il est bon de vérifier la cohérence des résultats obtenus, à savoir $\displaystyle\sum_{j=2}^{n} P(Y = j) = 1$ (le calcul est facile).

Nous vous conseillons de faire ce genre de vérification le plus souvent possible.

SF16.5 Utiliser la fonction génératrice

La fonction génératrice d'une variable aléatoire à valeurs entières est définie à partir de sa loi.

Mais compte tenu de l'unicité du développement en série entière, la connaissance de cette fonction peut permettre, à l'aide des coefficients, de trouver la loi.

Exemple

Une urne contient 4 boules numérotées 0,1,1,2. On effectue n tirages avec remise et on note S_n la somme des numéros tirés.

Déterminer la loi de S_n.

La fonction génératrice de S_1 est :

$$G_{S_1}(t) = P(S_1 = 0) + P(S_1 = 1)t + P(S_1 = 2)t^2 = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2 = \left(\frac{1+t}{2}\right)^2.$$

Les tirages étant avec remise, S_n est la somme de n variables aléatoires indépendantes de même loi que S_1. On sait que si des variables aléatoires sont indépendantes la fonction génératrice de leur somme est égale au produit des fonctions génératrices donc :

$$\forall\, t \in \mathbb{R},\ G_{S_n}(t) = \left[G_{S_1}(t)\right]^n = \left(\frac{1+t}{2}\right)^{2n}.$$

En développant cette expression avec la formule du binôme, on obtient :

$$\forall\, k \in [\![0\,;2n]\!],\ P(S_n = k) = \frac{1}{2^{2n}}\binom{2n}{k}.$$

Exemple

On réalise une succession de lancers d'une pièce où la probabilité d'obtenir pile est $p \in]0\,;1[$. Pour tout $n \geqslant 1$ on note X_n la variable aléatoire égale au nombre de lancers nécessaires pour obtenir le n^{e} pile.
On note ensuite $Y_1 = X_1$ et, pour $n \geqslant 1$, Y_n désigne le nombre de lancers supplémentaires nécessaires pour obtenir le n^{e} pile après l'obtention du $(n-1)^{\text{e}}$.
Déterminer la loi de Y_n et en déduire celle de X_n.

Une fois obtenu le $(n-1)^e$ pile, Y_n représente le nombre de lancers pour obtenir pile pour la 1re fois. Y_n suit donc la loi géométrique de paramètre p.
On a $X_n = Y_1 + \cdots + Y_n$; les lancers étant indépendants, les variables Y_n sont indépendantes. La fonction génératrice de leur somme est donc égale au produit des fonctions génératrices.
Puisque les Y_n suivent toutes la même loi géométrique de paramètre p, leur fonction génératrice est, d'après le cours, une série entière de rayon de convergence $\frac{1}{q}$ (avec $q = 1-p$), dont la somme vaut

$$\forall\, t \in \left]-\frac{1}{q}\,;\frac{1}{q}\right[,\ G_{Y_n}(t) = \frac{pt}{1-qt}\cdot$$

La fonction génératrice de X_n est donc donnée par :

$$\forall\, t \in \left]-\frac{1}{q}\,;\frac{1}{q}\right[,\ G_{X_n}(t) = \left(\frac{pt}{1-qt}\right)^n = p^n t^n \frac{1}{(1-qt)^n}\cdot$$

Pour obtenir la loi de X_n, il faut donc obtenir les coefficients du développement en série entière de cette fonction.
Or on a vu en **[SF13.12]** *page 522 que, pour tout $z \in \mathbb{C}$ tel que $|z| < 1$ et pour tout $p \in \mathbb{N}$, on a :*

$$\frac{1}{(1-z)^{p+1}} = \sum_{k=0}^{+\infty}\binom{k+p}{p} z^k.$$

Donc pour $|qt| < 1$ on a : $\dfrac{1}{(1-qt)^n} = \displaystyle\sum_{k=0}^{+\infty}\binom{k+n-1}{n-1} q^k t^k$, puis :

$$G_{X_n(t)} = p^n \sum_{k=0}^{+\infty}\binom{k+n-1}{n-1} q^k t^{n+k} = p^n \sum_{k=n}^{+\infty}\binom{k-1}{n-1} q^{k-n} t^k.$$

On a donc $X_n(\Omega) = [\![n, +\infty[\![$ *(ce qui est logique : pour obtenir* n *fois pile, il faut faire au moins* n *lancers) et :*

$$\forall\, k \geqslant n,\ P(X_n = k) = \binom{k-1}{n-1} p^n q^{k-n}\,.$$

On dit que X_n *suit une* loi de Pascal *de paramètres* (n,p).

✓ *On peut obtenir la loi de* X_n *par un calcul de probabilités direct : il suffit d'écrire que l'évènement* $(X_n = k)$ *est l'intersection des évènements « obtenir pile au* k^{e} *lancer » et « avoir obtenu* $n-1$ *pile lors des* $k-1$ *lancers précédents », ce dernier suivant un schéma binomial.*

■ Comment déterminer la loi d'un couple de variables aléatoires ?

SF16.6 Se ramener au calcul de la probabilité d'un évènement

Il s'agit ici de revenir simplement à la définition : après avoir déterminé les valeurs prises par deux variables aléatoires X et Y, on détermine la probabilité des évènements $(X = x) \cap (Y = y)$ pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$, ce qui peut se faire par les méthodes exposées dans le chapitre précédent et implique souvent l'usage des probabilités conditionnelles.

C'est ce que nous avons fait dans tous les exemples jusqu'à présent, nous ne détaillerons donc pas davantage.

SF16.7 Exploiter l'indépendance éventuelle de variables aléatoires

Si X et Y sont indépendantes, la loi du couple (X,Y) est immédiate puisque l'on a alors :

$$P\left(X = x \cap Y = y\right) = P(X = x) \times P(Y = y)\,.$$

On utilise fréquemment ce résultat pour déterminer la loi de variables aléatoires dépendant de X et de Y.

On peut aussi parfois simplifier les calculs en utilisant les résultats du cours sur les variables aléatoires indépendantes, en particulier ceux concernant les lois classiques.

Exemple

Soient X et Y deux variables aléatoires indépendantes, suivant toutes deux la loi géométrique de paramètre $p \in\,]0\,;1[$.
On pose $U = |X - Y|$ et $V = \min(X,Y)$.

1) Déterminer la loi du couple (U,V) .

2) En déduire les lois de U et de V.

3) U et V sont-elles indépendantes ?

1) *On remarque déjà que, puisque* X *et* Y *sont à valeurs dans* $\mathbb{N}^*$, U *est à valeurs dans* $\mathbb{N}$ *et* V *à valeurs dans* $\mathbb{N}^*$.
On va donc calculer les probabilités $P\left(U = i \cap V = j\right)$ *pour* $(i,j) \in \mathbb{N} \times \mathbb{N}^*$.

- *Le cas $i = 0$ est à traiter à part. En effet :*
$$P\left(U = 0 \cap V = j\right) = P\left(X = Y \cap X = j\right)$$
$$= P\left(X = j \cap Y = j\right) = \left(p(1-p)^{j-1}\right)^2 ,$$
car X et Y suivent des mêmes lois géométriques et sont indépendantes.
- *Pour $i \geqslant 1$, on aura*
$$(U = i) \cap (V = j) = \big((X - Y = i) \cup (Y - X = i)\big) \cap (V = j)$$
$$= \big[(X - Y = i) \cap (V = j)\big] \cup \big[(Y - X = i) \cap (V = j)\big]$$
$$= \big[(X - Y = i) \cap (Y = j)\big] \cup \big[(Y - X = i) \cap (X = j)\big]$$
et puisque les deux évènements ci-dessus sont incompatibles (car $i \neq 0$) :
$$P\left(U = i \cap V = j\right) = P\left(X = i + j \cap Y = j\right) + P\left(Y = i + j \cap X = j\right) .$$
Puisque X et Y sont indépendantes et de même loi géométrique, on obtient finalement :
$$P\left(U = i \cap V = j\right) = 2p(1-p)^{i+j-1}p(1-p)^{j-1} = 2p^2(1-p)^{i+2j-2}.$$

En résumé, pour tout $(i,j) \in \mathbb{N} \times \mathbb{N}^*$
$$P\left(U = i \cap V = j\right) = \begin{cases} p^2(1-p)^{2j-2} & \text{si } i = 0 \\ 2p^2(1-p)^{i+2j-2} & \text{sinon.} \end{cases}$$

2) *On obtient alors les lois marginales avec les formules habituelles, et en utilisant les formules sur les séries géométriques.*

- Loi de U
$$P(U = 0) = \sum_{j=1}^{+\infty} P\left(U = 0 \cap V = j\right) = p^2 \sum_{j=1}^{+\infty} [(1-p)^2]^{j-1} = \frac{p^2}{1-(1-p)^2}$$
$$= \frac{p}{2-p} ,$$
et pour tout $i \geqslant 1$:
$$P(U = i) = \sum_{j=1}^{+\infty} P\left(U = i \cap V = j\right) = 2p^2(1-p)^i \sum_{j=1}^{+\infty} [(1-p)^2]^{j-1}$$
$$= 2\frac{p(1-p)^i}{2-p} .$$
✓ *Nous vous laissons le soin de vérifier, toujours à l'aide des formules sur les séries géométriques, que l'on a bien* $\sum_{i=0}^{+\infty} P(U = i) = 1$.
- Loi de V

Pour tout $j \geqslant 1$ on a :
$$P(V = j) = \sum_{i=0}^{+\infty} P\left(U = i \cap V = j\right)$$
$$= P\left(U = 0 \cap V = j\right) + \sum_{i=1}^{+\infty} P\left(U = i \cap V = j\right)$$
$$= p^2(1-p)^{2j-2} + 2p^2(1-p)^{2j-2} \sum_{i=1}^{+\infty} (1-p)^i$$

$$P(V=j) = p^2(1-p)^{2j-2} + 2p^2(1-p)^{2j-2} \cdot \frac{1-p}{1-(1-p)}$$
$$= p(2-p)(1-p)^{2j-2}\,.$$

Là encore, on pourrait facilement vérifier que la somme de toutes ces probabilités est bien égale à 1.

3) *Il n'est alors pas bien difficile, en utilisant tous les calculs précédents, de vérifier que U et V sont bien indépendantes, c'est-à-dire que :*

$$P\left(U=i \cap V=j\right) = P(U=i)P(V=j)$$

pour tout $(i,j) \in \mathbb{N} \times \mathbb{N}^*$ (faire un calcul à part pour le cas $i=0$).

■ Comment étudier l'indépendance de deux variables aléatoires ?

SF16.8 Utiliser la définition

La notion d'indépendance, pour les variables aléatoires comme pour les évènements, est délicate et heurte souvent l'intuition première que l'on peut avoir (avant réflexion). Pour cette raison, la seule façon de démontrer que deux variables aléatoires sont indépendantes est de revenir à la définition et de vérifier si l'on a, ou non, une relation de la forme :

$$P\left(X=x_i \cap Y=y_j\right) = P(X=x_i) \times P(Y=y_j)\,.$$

Exemple

On suppose que le nombre N de colis expédiés chaque jour par une entreprise suit la loi de Poisson de paramètre λ. Les envois sont supposés indépendants.
La probabilité que chaque colis a d'arriver détérioré est égale à $p \in [0\,;1]$.
Pour un jour donné, on note X le nombre de colis qui sont arrivés détériorés, et Y le nombre de ceux arrivés en bon état.

1) Déterminer les lois de X et Y.

2) X et Y sont-elles indépendantes ?

1) *On remarque déjà que X et Y sont à valeurs dans $\mathbb{N}$.*

En utilisant le système complet d'évènements $\{(N=n),\, n \in \mathbb{N}\}$, la formule des probabilités totales donne :

$$\forall\, k \in \mathbb{N},\ P(X=k) = \sum_{n=0}^{+\infty} P\left(X=k \cap N=n\right) = \sum_{n=k}^{+\infty} P\left(X=k \cap N=n\right)$$

puisque l'évènement $(X=k) \cap (N=n)$ est impossible si $k>n$ (on ne peut pas avoir plus de colis détériorés que d'expédiés !).

D'autre part, si l'on sait que n colis ont été expédiés, le nombre de ceux qui arriveront détériorés suit la loi binomiale de paramètres (n,p) donc :

$$\forall\, n \in \mathbb{N},\ \forall\, k \in [\![0\,;n]\!],\ P\left(X=k \mid N=n\right) = \binom{n}{k} p^k (1-p)^{n-k}\,.$$

On en déduit, pour tout $k \in \mathbb{N}$:

$$\begin{aligned} P(X=k) &= \sum_{n=k}^{+\infty} P\left(X=k \mid N=n\right) P(N=n) \\ &= \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} \\ &= \mathrm{e}^{-\lambda} \frac{\lambda^k p^k}{k!} \sum_{n=k}^{+\infty} (1-p)^{n-k} \frac{\lambda^{n-k}}{(n-k)!} \\ &= \mathrm{e}^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{i=0}^{+\infty} \frac{(\lambda(1-p))^i}{i!} = \mathrm{e}^{-\lambda} \frac{(\lambda p)^k}{k!} \mathrm{e}^{\lambda(1-p)} = \mathrm{e}^{-\lambda p} \frac{(\lambda p)^k}{k!} . \end{aligned}$$

Donc X suit la loi de Poisson de paramètre λp, et de la même façon, Y suit la loi de Poisson de paramètre $\lambda(1-p)$.

2) *Pour i et j entiers, nous allons comparer $P\left(X=i \cap Y=j\right)$ avec $P(X=i)P(Y=j)$.*

En utilisant toujours le système complet d'évènements $\{(N=n),\, n \in \mathbb{N}\}$, on a :

$$P\left(X=i \cap Y=j\right) = \sum_{n=0}^{+\infty} P\big(X=i \cap Y=j \cap N=n\big) .$$

Or $N = X+Y$ donc si $n \neq i+j$ l'évènement $(X=i) \cap (Y=j) \cap (N=n)$ est impossible. Donc il ne reste qu'un terme dans la somme précédente et

$$\begin{aligned} P\left(X=i \cap Y=j\right) &= P\big(X=i \cap Y=j \cap N=i+j\big) \\ &= P\left(X=i \cap N=i+j\right) \end{aligned}$$

(on pouvait s'en douter...).
Les calculs sont semblables à ceux déjà faits ; on trouve :

$$\begin{aligned} P\left(X=i \cap N=i+j\right) &= P\left(X=i \mid N=i+j\right) P(N=i+j) \\ &= \left[\mathrm{e}^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!}\right] \left[\binom{i+j}{i} p^i (1-p)^j\right] = \mathrm{e}^{-\lambda} \lambda^{i+j} \frac{p^i (1-p)^j}{i!\, j!} . \end{aligned}$$

On constate donc que :

$$\forall\, (i,j) \in \mathbb{N}^2,\ P\left(X=i \cap Y=j\right) = P(X=i)P(Y=j) ,$$

c'est-à-dire que les variables aléatoires X et Y sont indépendantes !

✓ *Ce résultat peut sembler heurter le bon sens : il devrait y avoir un rapport entre le nombre de colis détériorés et le nombre de colis intacts. Sauriez-vous expliquer ce paradoxe ?*

SF16.9 Utiliser des fonctions de variables aléatoires indépendantes

Il s'agit ici d'exploiter le théorème suivant : si $(X_i)_{i \in I}$ est une famille de variables aléatoires indépendantes, si l'ensemble d'indices I est partitionné en $I = I_1 \cup I_2$ avec $I_1 \cap I_2 = \emptyset$, si U est une variable aléatoire fonction des $\{X_{i_1},\, i_1 \in I_1\}$ et V une variable aléatoire fonction des $\{X_{i_2},\, i_2 \in I_2\}$, alors U et V sont indépendantes.

Exemple

On lance n fois une pièce ($n \geqslant 2$). Soit n_1 un entier tel que $1 \leqslant n_1 < n$.
On note X_1 le nombre de fois où « pile » est apparu lors des n_1 premiers lancers, et X_2 le nombre de fois où « face » est apparu lors des lancers suivants.
Montrer que X_1 et X_2 sont indépendantes.

Si l'on note Y_k la variable de Bernoulli égale à 1 si le k^e tirage a donné « pile » et à 0 sinon, on a

$$X_1 = \sum_{k=1}^{n_1} Y_k \quad \text{et} \quad X_2 = \sum_{k=n_1+1}^{n} (1 - Y_k)\,.$$

Comme les lancers sont indépendants, il en est de même des variables Y_k donc de X_1 et X_2 d'après le théorème précédent.

✓ En fait, nous utilisons constamment le résultat démontré ci-dessus, en le considérant comme intuitivement évident et sans le redémontrer.

SF16.10 Exhiber un couple mettant en défaut l'indépendance

Exemple

Soit $n \in \mathbb{N}^*$ et p, q, r trois réels strictement positifs tels que $p+q+r=1$.
Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que la loi du couple (X,Y) est donnée par :

$$\forall\,(k,\ell) \in \mathbb{N}^2,\ P\big((X,Y)=(k,\ell)\big) = \begin{cases} \dfrac{n!}{k!\ell![n-(k+\ell)]!} p^k q^\ell r^{n-(k+\ell)} & \text{si } k+\ell \leqslant n \\ 0 & \text{sinon}\,. \end{cases}$$

1) Déterminer les lois de X et Y.

2) X et Y sont-elles indépendantes ?

On remarque déjà que, compte tenu de la définition de la loi de (X,Y), les variables aléatoires X et Y sont à valeurs dans $[\![0\,;n]\!]$.

1) *Par la formule des probabilités totales, pour $k \in [\![0\,;n]\!]$ on a :*

$$P(X=k) = \sum_{\ell=0}^{n-k} \frac{n!}{k!\ell![n-(k+\ell)]!} p^k q^\ell r^{n-(k+\ell)}$$

$$= p^k \frac{n!}{k!(n-k)!} \sum_{\ell=0}^{n-k} \binom{n-k}{\ell} q^\ell r^{(n-k)-\ell} = p^k \binom{n}{k} (q+r)^{n-k}$$

à l'aide de la formule du binôme.
Puisque $q+r = 1-p$, cela signifie que X suit la loi binomiale $\mathcal{B}(n,p)$.
Par symétrie, Y suit la loi binomiale $\mathcal{B}(n,q)$.

2) *On a $P\big((X,Y)=(0,0)\big) = r^n$ et :*

$$P(X=0) = (1-p)^n \quad , \quad P(Y=0) = (1-q)^n\,.$$

Puisque $r \neq (1-p)(1-q)$ (car $r = 1-p-q$ et $p,\,q>0$), on a

$$P\left(X=0 \cap Y=0\right) \neq P(X=0)P(Y=0)\,,$$

donc X et Y ne sont pas indépendantes.

SF16.11 Utiliser la covariance

Si deux variables aléatoires X et Y sont indépendantes, leur covariance est nulle, mais la réciproque est fausse.
On peut donc seulement en déduire que, si $\operatorname{cov}(X,Y) \neq 0$, alors X et Y ne sont pas indépendantes.

Exemple

On poursuit l'exemple précédent.

3) Calculer $E(XY)$ en utilisant la formule :

$$\forall\,(x,y) \in \mathbb{R}^2,\ \sum_{k=1}^{n} k\binom{n}{k} x^{k-1} y^{n-k} = n(x+y)^{n-1}\,.$$

4) En déduire la covariance de X et Y. Que peut-on en conclure ?

3) *La formule donnée par l'énoncé s'obtient facilement en dérivant par rapport à x la formule du binôme.*
En utilisant le théorème de transfert, *on a :*

$$E(XY) = \sum_{(k,\ell)\in\mathbb{N}^2} k\ell P\big((X,Y)=(k,\ell)\big)\,,$$

donc par définition de la loi du couple (X,Y) :

$$\begin{aligned}
E(XY) &= \sum_{k+\ell\leqslant n} k\ell \frac{n!}{k!\ell![n-(k+\ell)]!} p^k q^\ell r^{n-(k+\ell)} \\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{n-k} k\ell \frac{n!}{k!\ell![n-(k+\ell)]!} p^k q^\ell r^{n-(k+\ell)} \\
&= \sum_{k=1}^{n} kp^k \frac{n!}{k!(n-k)!}\, q\left(\sum_{\ell=1}^{n-k} \ell\binom{n-k}{\ell} q^{\ell-1} r^{(n-k)-\ell}\right) \\
&= \sum_{k=1}^{n} kp^k \frac{n!}{k!(n-k)!}\, q(n-k)(q+r)^{n-k-1} \\
&= \sum_{k=1}^{n-1} kp^k q(n-k)\binom{n}{k}(q+r)^{n-k-1}\,,
\end{aligned}$$

en ayant utilisé la formule de l'énoncé.
Ensuite, en utilisant la propriété : $(n-k)\binom{n}{k} = n\binom{n-1}{k}$, *on obtient :*

$$E(XY) = nqp\sum_{k=1}^{n-1} k\binom{n-1}{k} p^{k-1}(1-p)^{n-k-1} = nqp(n-1)[p+(1-p]^{n-1},$$

à l'aide encore de la formule de l'énoncé. Finalement : $E(XY) = n(n-1)pq$.

4) *Les variables X et Y suivant des lois binomiales de paramètres (n,p) et (n,q) respectivement, on a directement :*

$$\operatorname{cov}(X,Y) = E(XY) - E(X)E(Y) = n(n-1)pq - (np)(nq) = -npq\,.$$

Puisque $\operatorname{cov}(X,Y) \neq 0$, *$X$ et Y ne sont pas indépendantes.*

■ Comment étudier et déterminer l'espérance d'une variable aléatoire ?

SF16.12 Utiliser une formule liée à une loi classique

Si X suit une loi classique, il suffit d'appliquer directement la formule du cours pour obtenir son espérance.
Puisque les lois classiques apparaissent au détour de nombreux exercices, cette méthode est moins anodine qu'il n'y paraît. Nous l'avons d'ailleurs déjà utilisée plusieurs fois, et nous aurons encore l'occasion de le faire.

SF16.13 Utiliser la définition et appliquer la formule de base

On est ramené par la définition au calcul d'une somme finie ou de la somme d'une série ; dans ce dernier cas, il ne faut pas oublier de prouver la convergence.
Il n'est pas toujours facile de calculer ces sommes ; mais elles font souvent appel aux mêmes techniques, qui sont donc à bien connaître : sommes faisant intervenir les coefficients binomiaux (voir les rappels en **[S15.1]**), séries géométriques, séries entières (voir le thème 12).
Nous avons déjà eu l'occasion de faire ces calculs lors de nombreux exemples.

Exemple

On considère une variable aléatoire X à valeurs entières et une loi de probabilité P telles que :

$$\forall n \in \mathbb{N},\ P(X=n) = \frac{1}{(n+1)(n+2)}\cdot$$

Que peut-on dire de $E(X)$?

On vérifie qu'on définit bien ainsi une variable aléatoire et une probabilité car $P(X=n) \geqslant 0$ *pour tout* n *et* $\sum_{n=0}^{+\infty} P(X=n) = 1$. *En effet :*

$$\sum_{n=0}^{+\infty} \frac{1}{(n+1)(n+2)} = \sum_{n=0}^{+\infty} \left(\frac{1}{n+1} - \frac{1}{n+2}\right) = 1$$

après télescopage.
La variable X *n'admet pas d'espérance puisque*

$$nP(X=n) = \frac{n}{(n+1)(n+2)} \underset{n\to+\infty}{\sim} \frac{1}{n},$$

et la série $\sum_{n\in\mathbb{N}} nP(X=n)$ *diverge.*

SF16.14 Utiliser le théorème de transfert

Ce théorème est très important : en effet, il évite de déterminer la loi d'une variable aléatoire en permettant de calculer directement son espérance.

Exemple

Soit X une variable aléatoire suivant la loi de Poisson de paramètre $\lambda > 0$.

Calculer $E\left(\frac{1}{1+X}\right)$.

Sous réserve de convergence (absolue), le théorème de transfert donne :

$$E\left(\frac{1}{1+X}\right) = \sum_{n=0}^{+\infty} \frac{1}{n+1} P(X=n) = \mathrm{e}^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{(n+1)!}\,.$$

Il s'agit d'une série à terme réels positifs qui est clairement convergente (par exemple par la règle de d'Alembert). Cela prouve l'existence de $E\left(\frac{1}{1+X}\right)$. *De plus, puisque* $\lambda \neq 0$ *on peut écrire :*

$$E\left(\frac{1}{1+X}\right) = \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^{n+1}}{(n+1)!} = \frac{\mathrm{e}^{-\lambda}}{\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^n}{n!} = \frac{\mathrm{e}^{-\lambda}}{\lambda}\left(\mathrm{e}^{\lambda} - 1\right)$$

d'où finalement :

$$E\left(\frac{1}{1+X}\right) = \frac{1-\mathrm{e}^{-\lambda}}{\lambda}\,.$$

SF16.15 Appliquer le résultat lié à la série $\sum P(X \geqslant n)$

Nous avons déjà vu que dans certains cas, pour une variable aléatoire à valeurs entières, on obtient plus rapidement $P(X \geqslant n)$ que la loi de X. Pour calculer l'espérance de X on a alors tout intérêt à utiliser la formule :

$$E(X) = \sum_{n=1}^{+\infty} P(X \geqslant n) = \sum_{n=0}^{+\infty} P(X > n)$$

(sous réserve de convergence).

Exemple

On lance indéfiniment une pièce de monnaie équilibrée, et on note X la variable aléatoire correspondant au rang n où l'on a obtenu pour la 1re fois la séquence (pile, pile, pile).

Calculer l'espérance de X.

Pour $k \geqslant 3$, *considérons l'évènement :*

A_k *: « on a obtenu pile aux lancers de rangs* $k-2$, $k-1$ *et* k *sans jamais avoir obtenu la séquence (pile, pile, pile) auparavant ».*

Exactement comme dans un exemple de **[SF15.7]** *page 572, on peut montrer que l'évènement « obtenir au moins une fois la séquence (pile, pile, pile) », c'est-à-dire l'évènement* $\bigcup_{k=3}^{+\infty} A_k$, *est quasi-certain.*

On a donc, puisque les évènements A_k *sont incompatibles deux à deux :*

$$\sum_{k=3}^{+\infty} P(A_k) = P\left(\bigcup_{k=3}^{+\infty} A_k\right) = 1\,.$$

D'autre part, pour tout $k \geqslant 0$ *l'évènement* A_{k+4} *peut s'écrire :*

$$A_{k+4} = (X > k) \cap F_{k+1} \cap P_{k+2} \cap P_{k+3} \cap P_{k+4}$$

où F_i (respectivement P_i) désigne l'obtention d'un face (respectivement pile) au i^e tirage.
Puisque l'évènement $(X > k)$ signifie que l'on n'a pas obtenu la séquence voulue aux k premiers tirages, il ne dépend que de ces k premiers tirages, donc les évènements précédents sont indépendants d'où :

$$P(A_{k+4}) = P(X > k)P(F_{k+1})P(P_{k+2})P(P_{k+3})P(P_{k+4}) = \frac{1}{16}P(X > k)\,.$$

On en déduit que :

$$E(X) = \sum_{k=0}^{+\infty} P(X > k) = 16\sum_{k=0}^{+\infty} P(A_{k+4}) = 16(1 - P(A_3)) = 16\left(1 - \frac{1}{8}\right)$$

soit : $E(X) = 14$. Il s'agit du nombre moyen de lancers à effectuer pour obtenir trois pile successifs.

SF16.16 Utiliser la linéarité de l'espérance

Cette propriété permet parfois d'obtenir l'espérance d'une variable aléatoire sans avoir à calculer sa loi, simplement en l'expriment comme combinaison linéaire de variables aléatoires connues.
Reprenons l'exemple déjà traité en **[SF16.5]** page 592.

Exemple

On réalise une succession de lancers d'une pièce où la probabilité d'obtenir pile est $p \in]0\,;1[$. Pour tout $n \geqslant 1$ on note X_n la variable aléatoire égale au nombre de lancers nécessaires pour obtenir le n^e pile.
On note ensuite $Y_1 = X_1$ et, pour $n \geqslant 1$, Y_n désigne le nombre de lancers supplémentaires nécessaires pour obtenir le n^e pile après l'obtention du $(n-1)^e$.
Déterminer l'espérance et la variance de Y_n et en déduire celles de X_n.

Nous avions vu que chaque Y_n suit la loi géométrique de paramètre p. Donc d'après le cours $E(Y_n) = \frac{1}{p}$.
Puisque $X_n = Y_1 + \ldots + Y_n$ on en déduit directement et sans plus de calcul :

$$E(X_n) = \frac{n}{p}\,.$$

On sait aussi que la variance de Y_n est donnée par : $V(Y_n) = \frac{1-p}{p^2}$. Les variables Y_n étant indépendantes, la variance de X_n est donnée par :

$$V(X_n) = \sum_{i=1}^{n} V(Y_i) = \frac{n(1-p)}{p^2}\,.$$

✓ Nous avions déterminé la loi de X_n (loi de Pascal), mais cette méthode pour calculer espérance et variance est bien plus rapide que le calcul en utilisant la définition.

SF16.17 Utiliser l'indépendance dans l'étude d'un produit

On peut parfois utiliser la propriété : si deux variables X et Y sont indépendantes dont l'espérance existe, alors $E(XY) = E(X)E(Y)$.

Exemple

Un individu joue avec une pièce truquée, pour laquelle la probabilité d'obtenir pile est $p \in]0\,;1[$, de la façon suivante.

- Il lance la pièce jusqu'à obtenir pile pour la première fois. On note N la variable aléatoire égale au nombre de lancers nécessaires.
- Si n lancers ont été nécessaires pour obtenir pour la première fois pile, alors il relance n fois sa pièce. On appelle alors X le nombre de pile obtenu au cours de ces n lancers.

On admet la formule : $\forall x \in]-1\,;1[,\ \displaystyle\sum_{n=k}^{+\infty} \binom{n}{k} x^n = \frac{x^k}{(1-x)^{k+1}}$.

1) Déterminer la loi de N.

2) En déduire la loi de X.

3) On considère deux variables aléatoires B et G indépendantes, suivant respectivement la loi de Bernoulli $\mathcal{B}(1,p')$ et la loi géométrique $\mathcal{G}(p')$.
Montrer qu'on peut choisir p' tel que X a la même loi que la variable BG, et en déduire $E(X)$.

1) *La variable aléatoire N correspond au rang d'apparition pour la première fois de l'évènement « obtenir pile » , qui est de probabilité p, au cours d'une succession d'épreuves identiques. N suit donc la loi géométrique de paramètre p.*

2) *La variable X est à valeurs dans $\mathbb{N}$. La formule des probabilités totales avec le système complet d'évènements $\{(N=n),\ n \in \mathbb{N}^*\}$ donne :*

$$\forall k \in \mathbb{N},\ P(X=k) = \sum_{n=1}^{+\infty} P(X=k \mid N=n)\, P(N=n)\,.$$

On cherche donc à calculer $P(X=k \mid N=n)$.
Or, lorsque $N=n$, X compte le nombre de réalisations de l'évènement « obtenir pile » , qui est de probabilité p, au cours de n réalisations identiques d'une épreuve de Bernoulli. La loi conditionnelle de X sachant $(N=n)$ est donc la loi binomiale de paramètres n et p :

$$\forall k \in \mathbb{N},\ P(X=k \mid N=n) = \begin{cases} \binom{n}{k} p^k q^{n-k} & \text{si } k \leqslant n \\ 0 & \text{sinon} \end{cases}$$

en ayant noté $q = 1-p$.
On en déduit, en utilisant la formule donnée par l'énoncé (qui s'obtient simplement en dérivant k fois le développement en série entière de $\dfrac{1}{1-x}$, voir **[SF13.12]**, *page 522), pour $k \geqslant 1$:*

$$\begin{aligned} P(X=k) &= \sum_{n=k}^{+\infty} P(X=k \mid N=n)\, P(N=n) \\ &= \sum_{n=k}^{+\infty} pq^{n-1} \binom{n}{k} p^k q^{n-k} = p^{k+1} q^{-k-1} \sum_{n=k}^{+\infty} \binom{n}{k} (q^2)^n \\ &= \frac{p^{k+1}}{q^{k+1}} \cdot \frac{q^{2k}}{(1-q^2)^{k+1}} = \frac{p^{k+1} q^{k-1}}{(1-q)^{k+1}(1+q)^{k+1}} = \frac{q^{k-1}}{(1+q)^{k+1}} \end{aligned}$$

et pour $k = 0$:

$$P(X=0) = \sum_{n=1}^{+\infty} pq^{n-1}q^n = \frac{p}{q}\sum_{n=1}^{+\infty}(q^2)^n = \frac{pq^2}{q(1-q^2)} = \frac{q}{1+q}\cdot$$

3) *La variable BG est à valeurs dans $\mathbb{N}$.*
On a : $P(BG=0) = P(B=0) = 1-p'$ et, pour $k \geqslant 1$:

$$\begin{aligned} P(BG=k) &= P\left(B=1 \cap G=k\right) = P(B=1)P(G=k) \\ &= p' \cdot (1-p')^{k-1}p' = p'^2(1-p')^{k-1}\,. \end{aligned}$$

On vérifie qu'il suffit de choisir $p' = 1 - \dfrac{q}{q+1} = \dfrac{1}{1+q}$ pour que la loi de BG soit la même que celle de X.
Comme X et BG ont la même loi, elles ont la même espérance. On a donc $E(X) = E(BG)$. Mais B et G sont indépendantes, donc $E(BG) = E(B)E(G) = p' \times \dfrac{1}{p'} = 1$ et finalement : $E(X) = 1$.

SF16.18 Utiliser la fonction génératrice

Si X est une variable aléatoire dont la fonction génératrice G_X est dérivable à gauche en 1, alors son espérance existe et :

$$E(X) = G'_X(1^-)\,.$$

Donc si l'on a déjà calculé la fonction génératrice, il suffit d'appliquer cette formule pour obtenir l'espérance.
Reprenons un exemple déjà vu en **[SF16.5]** page 591 .

Exemple

Une urne contient 4 boules numérotées 0,1,1,2. On effectue n tirages avec remise et on note S_n la somme des numéros tirés.

Déterminer l'espérance de S_n.

Nous avions calculé : $\forall\, t \in \mathbb{R},\ G_{S_n}(t) = \left(\dfrac{1+t}{2}\right)^{2n}$.

G_{S_n} est dérivable sur $\mathbb{R}$ et : $\forall\, t \in \mathbb{R},\ G'_{S_n}(t) = n\left(\dfrac{1+t}{2}\right)^{2n-1}$.

D'où directement : $E(S_n) = G'_{S_n}(1) = n$.

On pouvait aussi ici utiliser la linéarité de l'espérance, puisque S_n est somme de n variables aléatoires de même loi et d'espérance 1.

■ Comment étudier et déterminer la variance d'une variable aléatoire ?

SF16.19 Utiliser une formule liée à une loi classique

Si X suit une loi classique, il suffit d'appliquer directement la formule du cours pour obtenir sa variance.

SF16.20 Utiliser la formule de Kœnig-Huygens

On calcule rarement la variance d'une variable aléatoire X (si elle existe) en utilisant la définition :

$$V(X) = E\left(\left(X - E(X)\right)^2\right).$$

mais plutôt en utilisant la formule de Kœnig-Huygens :

$$V(X) = E(X^2) - E(X)^2.$$

L'espérance de X^2 peut se calculer à l'aide de la formule de transfert :

$$E(X^2) = \sum_{x \in X(\Omega)} x^2 P(X = x).$$

Exemple

X et Y sont deux variables aléatoires indépendantes, suivant toutes deux la loi uniforme sur $[\![0\,;n]\!]$.
On note Z la variable aléatoire : $Z = |X - Y|$.

Calculer la variance de Z.

Nous allons répondre à cette question en essayant d'utiliser au maximum toutes les propriétés du cours, afin de minimiser le nombre de calculs (en particulier, nous ne déterminerons pas la loi de Z).

- *On calcule l'espérance de Z en utilisant le théorème de transfert et le fait que X et Y sont indépendantes.*

$$\begin{aligned} E(Z) &= \sum_{(i,j)\in[\![0;n]\!]^2} |i-j|\, P\left(X = i \cap Y = j\right) \\ &= \sum_{(i,j)\in[\![0;n]\!]^2} |i-j|\, P(X=i)P(Y=j), \end{aligned}$$

et puisque $P(X = i) = P(Y = j) = \dfrac{1}{n+1}$ pour tous i, j :

$$E(Z) = \frac{1}{(n+1)^2} \sum_{(i,j)\in[\![0;n]\!]^2} |i-j| = \frac{2}{(n+1)^2} \sum_{i=0}^{n} \sum_{j=0}^{i-1} (i-j)$$

$$= \frac{2}{(n+1)^2} \sum_{i=0}^{n} \left(i^2 - \sum_{j=0}^{i-1} j \right) = \frac{2}{(n+1)^2} \sum_{i=0}^{n} \left(i^2 - \frac{i(i-1)}{2} \right)$$

$$= \frac{2}{(n+1)^2} \sum_{i=0}^{n} \left(\frac{i^2+i}{2} \right) = \frac{1}{(n+1)^2} \left(\frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2} \right)$$

en ayant utilisé les formules (à connaître) :

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Après simplification, on trouve : $E(Z) = \dfrac{n(n+2)}{3(n+1)}$.

- *Par linéarité de l'espérance :*

$$E(Z^2) = E\big((X-Y)^2\big) = E(X^2) + E(Y^2) - 2E(XY)\,.$$

Puisque X et Y ont la même loi et sont indépendantes :

$$E(Z^2) = 2E(X)^2 - 2[E(X)]^2 = 2V(X)\,.$$

Ensuite :

$$E(X^2) = \sum_{k=0}^{n} k^2 \cdot \frac{1}{n+1} = \frac{n(2n+1)}{6}\,,$$

et

$$E(X) = \sum_{k=0}^{n} k \cdot \frac{1}{n+1} = \frac{n}{2}$$

d'où

$$V(X) = \frac{n(2n+1)}{6} - \frac{n^2}{4} = \frac{n(n+2}{12}$$

et enfin : $V(Z) = E(Z^2) - [E(Z)]^2 = \dfrac{n(n+2)}{6} - \dfrac{n^2(n+2)^2}{9(n+1)^2}$

$$= \frac{n(n+2)(n^2+2n+3)}{18(n+1)^2}\,.$$

✓ *Une fois tous les calculs terminés, il est bon de vérifier que l'on a bien* $V(Z) \geqslant 0$ *!*

✓ *On ne pouvait pas ici utiliser directement les formules du cours donnant l'espérance et la variance d'une variable aléatoire suivant une loi uniforme, car ces formules sont données pour une variable à valeurs dans* $[\![1\,;n]\!]$ *et ici, X et Y sont à valeurs dans* $[\![0\,;n]\!]$.

SF16.21 Utiliser la variance d'une somme de variables indépendantes

Si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes, alors

$$V(X_1 + \ldots + X_n) = V(X_1) + \ldots + V(X_n)\,.$$

Cette propriété peut permettre de calculer la variance d'une variable aléatoire sans avoir à déterminer sa loi.

Exemple

Une urne contient N boules numérotées de 1 à N. On effectue n tirages avec remise et on appelle S_n la somme des numéros obtenus.

Déterminer l'espérance et la variance de S_n.

Pour tout $i \in [\![1\,;n]\!]$, notons X_i la variable aléatoire égale au numéro sorti lors du i^e tirage. Ainsi : $S_n = \sum_{i=1}^{n} X_i$.

Or les X_i sont indépendantes, et chaque X_i suit la loi uniforme sur $[\![1\,;n]\!]$ donc a pour espérance $\dfrac{N+1}{2}$ et pour variance $\dfrac{N^2-1}{12}$.

On en déduit directement :

$$E(S_n) = n\frac{N+1}{2} \quad \text{et} \quad V(X) = n\frac{N^2-1}{12}\,.$$

SF16.22 Utiliser la fonction génératrice

Si X est une variable aléatoire dont la fonction génératrice G_X est deux fois dérivable à gauche en 1, alors sa variance existe et :

$$E(X) = G'_X(1^-) \quad , \quad E(X^2) = G'_X(1^-) + G''_X(1^-)\,,$$

donc si l'on connaît G_X on peut facilement obtenir $V(X)$ par ces formules et la formule de Kœnig-Huygens.

■ Comment calculer une covariance ?

SF16.23 Appliquer la formule de type Kœnig-Huygens

On ne calcule jamais la covariance de deux variables aléatoires X et Y en utilisant la définition :

$$\text{cov}(X,Y) = E\big[(X - E(X))(Y - E(Y))\big]$$

mais plutôt en utilisant la formule :

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y)\,.$$

Dans cette dernière formule, si l'on connaît la loi du couple (X,Y), on calcule l'espérance du produit à l'aide du théorème de transfert.
Nous avons déjà vu un exemple de ce calcul en **[SF16.11]** page 598.

SF16.24 Faire le lien avec la variance d'une somme

Par la formule :

$$V(X+Y) = V(X) + V(Y) + 2\text{cov}\,(X,Y)\,,$$

on peut obtenir la covariance de X et Y connaissant la variance de $X+Y$, ou bien calculer la variance de la somme connaissant la covariance.

Exemple

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes, suivant la loi de Bernoulli de paramètre $p \in]0\,;1[$.
Pour tout $n \in \mathbb{N}^*$ on pose : $Y_n = X_n X_{n+1}$ et $S_n = Y_1 + \cdots + Y_n$.

Calculer l'espérance et la variance de S_n

Nous ne chercherons pas la loi de S_n mais seulement celle des Y_n.

- *La variable Y_n est à valeurs dans $\{0,1\}$ et*

  $$P(Y_n = 1) = P\left(X_n = 1 \cap X_{n+1} = 1\right) = P(X_n = 1)P(X_{n+1} = 1) = p^2$$

  puisque X_n et X_{n+1} sont indépendantes.
  Y_n suit donc la loi de Bernoulli de paramètre p^2, donc $E(Y_n) = p^2$ et $V(Y_n) = p^2(1-p^2)$.
- *On en déduit immédiatement que :*

  $$E(S_n) = \sum_{i=1}^{n} E(Y_i) = np^2.$$

- *On a ensuite :*

$$V(S_n) = \sum_{i=1}^{n} V(Y_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \operatorname{cov}(Y_i, Y_j)\,.$$

Or si $|n-m| \geqslant 2$ les variables aléatoires Y_n et Y_m sont indépendantes puisque Y_n ne dépend que de X_n et X_{n+1} et Y_m que de X_m et X_{m+1} et que les deux ensembles $\{X_n, X_{n+1}\}$ et $\{X_m, X_{m+1}\}$ sont disjoints.
On aura donc, pour $i < j$ et $j \neq i+1$, $\operatorname{cov}(Y_i, Y_j) = 0$.
Il ne reste donc plus qu'à calculer, pour $i \in [\![1\,;n-1]\!]$:

$$\operatorname{cov}(Y_i, Y_{i+1}) = E(Y_i Y_{i+1}) - E(Y_i)E(Y_{i+1})\,.$$

Or $Y_i Y_{i+1}$ est une variable aléatoire à valeurs dans $\{0,1\}$ et pour $i \leqslant n-2$:

$$Y_i Y_{i+1} = 1 \iff Y_i = Y_{i+1} = 1 \iff X_i = X_{i+1} = X_{i+2} = 1$$

donc $P(Y_i Y_{i+1} = 1) = p^3$; $Y_i Y_{i+1}$ suit donc la loi de Bernoulli de paramètre p^3 d'où :

$$\operatorname{cov}(Y_i, Y_{i+1}) = p^3 - (p^2)^2 = p^3(1-p)\,,$$

et finalement :

$$V(X_n) = np^2(1-p^2) + 2(n-1)p^3(1-p) = p^2(1-p)(3np + n - 2p)\,.$$

■ Comment encadrer une probabilité ?

SF16.25 Utiliser l'inégalité de Markov

L'inégalité de Markov est vraiment trop grossière pour pouvoir estimer correctement une probabilité.
Elle n'est vraiment utile qu'en théorie ; elle a d'ailleurs servi dans le cours pour démontrer l'inégalité de Bienaymé-Tchebychev qui, elle, est plus précise puisqu'elle fait aussi intervenir la variance. En voici un autre exemple d'application.

Exemple

Soit X une variable aléatoire telle que $E(|X|) = 0$ Montrer que X est quasi certainement nulle.

D'après l'inégalité de Markov :

$$\forall\, a > 0,\ P(|X| \geqslant a) \leqslant \frac{E(|X|)}{a} = 0\,.$$

Or :

$$(|X| > 0) = \bigcup_{n=1}^{+\infty} \left(|X| \geqslant \frac{1}{n}\right)$$

d'où :

$$P(|X| > 0) \leqslant \sum_{n=1}^{+\infty} P\left(|X| \geqslant \frac{1}{n}\right) = 0$$

donc $P(|X| > 0) = 0$ et par suite $P(|X| = 0) = 1$: l'évènement $(X = 0)$ est quasi certain.

SF16.26 Utiliser l'inégalité de Bienaymé-Tchebychev

Exemple

On lance n fois une pièce équilibrée, et on note X le nombre de fois où le lancer a donné pile.
Pour tout $\lambda > 0$, donner un majorant de $P\left(\left|X - \frac{n}{2}\right| \geqslant \lambda n\right)$.

X suit la loi binomiale $\mathcal{B}(n,\frac{1}{2})$ donc directement d'après le cours :

$$E(X) = \frac{n}{2} \quad \text{et} \quad V(X) = \frac{n}{4}\cdot$$

L'inégalité de Bienaymé-Tchebychev donne alors, pour tout $\lambda > 0$:

$$P\big(|X - E(X)| \geqslant \lambda n\big) \leqslant \frac{V(X)}{\lambda^2 n^2}$$

soit :

$$P\left(\left|X - \frac{n}{2}\right| \geqslant \lambda n\right) \leqslant \frac{1}{4\lambda^2 n^2}\cdot$$

■ Comment utiliser les résultats asymptotiques concernant les variables aléatoires ?

SF16.27 Utiliser l'approximation de la loi binomiale par la loi de Poisson

Si $(p_n)_{n\in\mathbb{N}^*}$ est une suite de réels de l'intervalle $]0\,;1[$ telle que $\lim\limits_{n\to+\infty} np_n = \lambda > 0$, alors la loi binomiale de paramètres (n,p_n) tend vers la loi de Poisson de paramètre λ.
Ce résultat permet de comprendre pourquoi la loi de Poisson est si souvent utilisée lorsqu'on cherche à estimer le nombre de fois où un évènement se réalise pendant un temps donné.
Dans la pratique, on admet que cette approximation est satisfaisante dès que $n \geqslant 30$, $p \leqslant 0{,}1$ avec $np \leqslant 10$. Lorsque les calculs étaient faits « à la main », cela permettait un gain de temps notable. Pour nous, cela sert surtout à simplifier les calculs théoriques car il est plus facile de travailler avec la fonction exp qu'avec des coefficients binomiaux.

Exemple

Un laboratoire met sur le marché un médicament qui a une très faible probabilité p d'avoir des effets indésirables majeurs.
Lors d'un test sur 1 000 personnes, aucun effet indésirable n'a été relevé. Le laboratoire affirme alors que la probabilité que son médicament soit inoffensif est supérieure à 99%. Est-ce sérieux ?

Si X est le nombre de personnes qui sont victimes d'un effet indésirable dans une population d'effectif n, X suit la loi binomiale $\mathcal{B}(n,p)$. p étant ici très petit et n très grand, on peut approcher la loi de X par la loi de Poisson $\mathcal{P}(np)$.

Ainsi, $P(X=0)=\mathrm{e}^{-np}$.
Le laboratoire considère donc $P(X=0)\geqslant 99\%$. *Cela correspond à une valeur de* p *telle que*

$$\mathrm{e}^{-1\,000p}\geqslant 0{,}99$$

soit

$$p\leqslant -\frac{1}{1\,000}\ln(0{,}99)\approx 10^{-5}\,.$$

Mais avec $p=10^{-5}$, *si le médicament est administré à* $10\,000\,000$ *personnes, l'espérance (si l'on ose dire) de la variable* X *est égale à* $10\,000\,000p$, *c'est-à-dire à* 100. 100 *victimes, cela fait beaucoup !*

> ✓ *Il s'agit ici d'une approche très simpliste des problèmes qui se posent en statistiques : comment à partir d'observations déduire des conclusions générales ? Notre exemple n'avait comme but que de montrer l'intérêt d'approcher la loi binomiale par la loi de Poisson pour simplifier les calculs.*

SF16.28 Appliquer la loi faible des grands nombres

La loi faible des grands nombres peut servir à mesurer l'écart entre une fréquence observée et une fréquence théorique.

Exemple

On a lancé 5 000 fois une pièce de monnaie et l'on a obtenu 2600 fois pile.
Déterminer un intervalle tel que la probabilité p d'obtenir pile appartienne à cet intervalle avec une probabilité d'au moins 95%.

On applique ici la loi faible des grands nombres à la suite (X_n) *des variables de Bernoulli, telles que que* $X_n=1$ *si on a obtenu pile au* n^e *lancer, et* $X_n=0$ *sinon.*

La variable $\overline{X_n}=\dfrac{\sum_{i=1}^{n}X_i}{n}$ *représente la fréquence d'obtention de pile au cours de* n *lancers.*

Par linéarité, on a : $E(\overline{X_n})=\dfrac{1}{n}\sum_{i=1}^{n}E(X_i)=\dfrac{1}{n}\cdot np=p$.

Puis, les variables X_i *étant indépendantes de même variance* $p(1-p)$ *on a :*

$$\begin{aligned}V(\overline{X_n})&=V\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}V\left(\sum_{i=1}^{n}X_i\right)\\&=\frac{1}{n^2}\sum_{i=1}^{n}V(X_i)=\frac{1}{n^2}\cdot np(1-p)=\frac{p(1-p)}{n}\end{aligned}$$

L'inégalité de Bienaymé-Tchebychev donne alors :

$$\forall\,\varepsilon>0,\ P\left(\left|\overline{X_n}-p\right|\geqslant\varepsilon\right)\leqslant\frac{p(1-p)}{n\varepsilon^2}\leqslant\frac{1}{4n\varepsilon^2}\,,$$

car il est facile de vérifier que : $\forall\,x\in[0\,;1],\ x(1-x)\leqslant\dfrac{1}{4}$.

Cela peut aussi s'écrire : $P\left(\left|\overline{X_n} - p\right| < \varepsilon\right) \geqslant 1 - \frac{1}{4n\varepsilon^2}$ *. Puisque l'on cherche* ε *tel que* $P\left(\left|\overline{X_n} - p\right| < \varepsilon\right) \geqslant 0{,}95$*, il suffit de le choisir tel que :*

$$1 - \frac{1}{4n\varepsilon^2} \geqslant 0{,}95 \quad \textit{soit} \quad n\varepsilon^2 \geqslant 5\,.$$

Ici $n = 5\,000$ *et l'on prendra donc* $\varepsilon \approx 0{,}032$*.*

Puisque $\overline{X_n} = \dfrac{2\,600}{5\,000} = 0{,}52$*, on aura donc* $p \in \,]0{,}488\,;\,0{,}552[$ *avec une probabilité supérieure à 95%.*

✓ *On ne s'est pas contenté d'appliquer directement la loi faible des grands nombres, qui donne seulement une limite, mais on en a refait la démonstration à partir de l'inégalité de Bienaymé-Tchebychev.*

À vous de jouer !

Exercice 16.1 **SF15.8** **SF16.13** **SF16.20**

Un sauteur en hauteur tente de franchir les hauteurs successives $1, 2, \ldots, n, \ldots$
On suppose les sauts indépendants les uns des autres, et l'on suppose que la probabilité de réussite à la hauteur n est égale à $\dfrac{1}{n}$.
Le sauteur est éliminé à son premier échec.
On note X la variable aléatoire égale au numéro du dernier saut réussi.

1) Trouver la loi de X et vérifier que : $\displaystyle\sum_{n=1}^{+\infty} P(X = n) = 1$.

2) Calculer l'espérance et la variance de X.

Exercice 16.2 **SF16.1** **SF16.16** **SF16.19** **SF16.21** **SF16.26**

Un fabricant de chocolat met dans chacune de ses tablettes une image faisant partie d'une collection de n images distinctes. On suppose que chaque tablette contient une des n images avec la même probabilité, et que les images sont indépendantes d'une tablette à l'autre.

On note X_n la variable aléatoire égale au nombre de tablettes qu'il faut acheter pour réunir la collection complète.

1) Calculer l'espérance de X_n, et en donner un équivalent quand $n \to +\infty$.

2) Calculer la variance de X_n. Montrer qu'il existe une constante $C > 0$ telle que $V(X_n) \leqslant Cn^2$ et en déduire :

$$\forall\, \varepsilon > 0,\ \lim_{n\to+\infty} P\left(\left|\frac{X_n}{n \ln n} - 1\right| \geqslant \varepsilon\right) = 0\,.$$

Exercice 16.3 SF16.1 SF16.6 SF16.7

Le nombre d'œufs pondus par une poule en 1 mois est une variable aléatoire qui suit la loi de Poisson de paramètre α.
Le nombre d'œufs pondus par une cane en 1 mois est une variable aléatoire qui suit la loi de Poisson de paramètre β.
On suppose que ces deux variables sont indépendantes.

Si le fermier ramasse n œufs en 1 mois, quelle est la probabilité que k d'entre eux soient des œufs de poule ?

Exercice 16.4 SF16.2 SF16.13 SF16.6 SF16.10

Soit n un entier supérieur ou égal à 2. On considère une urne contenant une boule numérotée « 1 », deux boules numérotées « 2 », ..., n boules numérotées « n ».

1) On tire une boule de cette urne et on note X le numéro obtenu. Déterminer la loi de X et calculer $E(X)$.

2) On effectue dans cette urne deux tirages successifs sans remise. On note toujours X le numéro de la première boule obtenue et Y le numéro de la deuxième boule.

a) Déterminer la loi du couple (X,Y).

b) En déduire la loi de Y.

c) Les variables aléatoires X et Y sont-elles indépendantes ?

Exercice 16.5 SF15.2 SF16.4 SF16.10 SF16.14

Soit (X,Y) un couple de variables aléatoires à valeurs dans $\mathbb{N}^2$, de loi conjointe :

$$\forall\,(i,j)\in\mathbb{N}^2,\ P\left(X=i\cap Y=j\right)=\frac{(i+j)\lambda^{i+j}}{\mathrm{e}\,i!j!}\,,$$

où λ est un paramètre réel >0.

1) Que vaut λ ?

2) Déterminer les lois de X et de Y.

3) X et Y sont-elle indépendantes ?

4) Calculer $E\left(2^{X+Y}\right)$.

Exercice 16.6 SF15.3 SF16.2 SF16.13 SF16.20 SF16.24

Soient N et a deux entiers, $N\geqslant 2$, $a\geqslant 1$.
On dispose d'une urne contenant aN boules, de N couleurs différentes, à raison de a boules par couleur.
On tire n boules simultanément, et on note X_i le nombre de boules de la couleur n° i obtenues.

1) Déterminer la loi de X_i pour $1 \leqslant i \leqslant N$, puis les valeurs de $E(X_i)$ et de $V(X_i)$.

On rappelle la formule de Vandermonde :

$$\forall\,(n,m,k) \in \mathbb{N}^3,\ \sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}.$$

2) En utilisant l'égalité :

$$X_1 + X_2 + \cdots + X_N = n\,,$$

déterminer $\operatorname{cov}(X_i,X_j)$ et le coefficient de corrélation $\rho(X_i,X_j)$ pour $i \neq j$.

Exercice 16.7 — SF16.5

On considère deux dés truqués ; pour $1 \leqslant i \leqslant 6$, la probabilité d'apparition du chiffre i est p_i pour le premier et q_i pour le second, avec p_i et q_i strictement positifs tels que $\sum_{i=1}^{6} p_i = \sum_{i=1}^{6} q_i = 1$.
On lance les deux dés et on désigne par X_1 (respectivement X_2) la variable aléatoire égale au point amené par le premier dé (respectivement le second).

En considérant les fonctions génératrices de X_1 et X_2, montrer qu'il n'est pas possible de choisir les p_i et les q_i de sorte que $X_1 + X_2$ suive la loi uniforme.

Exercice 16.8 — SF15.10 SF16.5

On désigne par N le nombre de champignons ramassés par un individu durant une période donnée. On suppose que N est une variable aléatoire à valeurs dans $\mathbb{N}^*$ de fonction génératrice G_N.
On suppose de plus que la probabilité pour qu'un champignon cueilli soit comestible est égale à p.

Montrer que la probabilité que *tous* les champignons ramassés soient comestibles est égale à $G_N(p)$.

Exercice 16.9 — SF15.10 SF16.5

Soit N et X_1, $X_2, \ldots$ des variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. On suppose que les variables X_1, $X_2, \ldots$ suivent toutes une même loi de fonction génératrice G_X et on pose :

$$S = \sum_{k=1}^{N} X_k\,.$$

1) Montrer que : $G_S(t) = G_N\left(G_X(t)\right)$ pour $|t| \leqslant 1$

2) On suppose que les variables admettent toutes une espérance. Établir l'identité de Wald :

$$E(S) = E(N)E(X_1)\,.$$

Thème 17 - Équations différentielles linéaires

Seules figurent désormais au programme les équations différentielles *linéaires*. Le seul résultat important concernant ces équations est le théorème de Cauchy-Lipschitz, dont il faut bien connaître les hypothèses. Il faut aussi connaître certaines techniques de calcul, mais celles-ci sont en nombre assez limité.

■ Comment résoudre une équation différentielle scalaire linéaire d'ordre 1 ?

La réponse est pour une fois très simple : il suffit d'appliquer directement les résultats du cours ! Mais attention : tous ces résultats, en particulier le théorème de Cauchy-Lipschitz qui permet d'obtenir la dimension de l'espace des solutions, s'appliquent à une équation *réduite*, c'est-à-dire où le coefficient de y' est égal à 1. Lorsque le coefficient de y' peut s'annuler, il faut appliquer ces résultats sur chacun des intervalles où il ne s'annule pas, puis éventuellement raccorder les solutions.

SF17.1 Appliquer les résultats du cours

Exemple

Chercher les solutions de l'équation différentielle

$$(E) \qquad x(x^2-1)y' + 2y = x^2.$$

Existe-t-il une solution définie sur $\mathbb{R}$?

- *Puisque le coefficient de y' s'annule pour $x \in \{-1,0,1\}$, on résout l'équation séparément sur chacun des intervalles :*

$$I_1 =]-\infty\,;-1[\,, \quad I_2 =]-1\,;0[\,, \quad I_3 =]0\,;1[\quad \textit{et} \quad I_4 =]1\,;+\infty[.$$

- *L'équation homogène associée à (E) est :*

$$(E_0) \quad y' = a(x)y \quad \textit{avec} \quad a(x) = \frac{-2}{x(x^2-1)}\cdot$$

Une primitive de a sur chacun des intervalles considérés est :

$$A(x) = \int \frac{-2}{x(x^2-1)}\,\mathrm{d}x = 2\int\left(\frac{1}{x} - \frac{x}{x^2-1}\right)\mathrm{d}x = 2\ln|x| - \ln\left|x^2-1\right|\,.$$

✓ *N'oubliez pas les valeurs absolues !*

Les solutions de E_0 sont donc d'après le cours les fonctions de la forme :

$$x \mapsto \lambda \exp A(x) = \lambda\frac{x^2}{x^2-1}\,,$$

la constante λ dépendant de l'intervalle.

- *On finit de résoudre l'équation par la méthode de variation de la constante. Sur chacun des intervalles, l'équation (E) s'écrit sous forme résolue :*

$$y' - a(x)y = \frac{x}{x^2-1}\cdot$$

On cherche la solution générale de (E) sous la forme $y(x) = \dfrac{x^2}{x^2-1}\varphi(x)$. On obtient :

$$y \text{ solution de } (E) \iff \frac{x^2}{x^2-1}\varphi'(x) = \frac{x}{x^2-1} \iff \varphi'(x) = \frac{1}{x}\cdot$$

✓ *Nous n'avons fait ici aucun calcul compliqué de dérivée ! Mais simplement appliqué directement les résultats du cours : on savait qu'il ne resterait plus que les termes en $\varphi'(x)$. Il faut cependant faire attention : ce résultat s'applique à l'équation mise sous forme résolue.*

On a donc $\varphi(x) = \ln|x|$ et la solution générale de (E) sur chaque intervalle I_k est de la forme :

$$y = \frac{x^2}{x^2-1}\left(\ln|x| + \lambda_k\right), \quad k \in [\![1\,;4]\!]\,.$$

- Raccordement des solutions

Déjà, toutes les solutions se raccordent bien en 0 puisque :

$$\forall\, k \in [\![1\,;4]\!],\ \lim_{x\to 0} \frac{x^2}{x^2-1}\left(\ln|x| + \lambda_k\right) = 0\,.$$

On connaît la limite $\lim\limits_{x\to 1} \dfrac{\ln x}{x-1} = 1$ d'où :

$$\lim_{x\to 1} \frac{x^2}{x^2-1}\ln|x| = \frac{1}{2} \quad \text{et} \quad \lim_{x\to -1} \frac{x^2}{x^2-1}\ln|x| = \frac{1}{2}\cdot$$

Donc pour que les solutions se raccordent bien en 1, il faut $\lambda_3 = \lambda_4 = 0$, et pour qu'elles se raccordent en -1 il faut $\lambda_1 = \lambda_2 = 0$.

La seule solution possible définie sur $\mathbb{R}$ est donc :

$$f\colon x \mapsto \frac{x^2\ln|x|}{x^2-1}\cdot$$

Pour que cette fonction convienne, il faut en plus que ses dérivées existent et se raccordent bien en -1, 0 et 1.

Or pour $x \in \bigcup\limits_{i=1}^{4} I_k$, on a (sans calcul, en utilisant directement l'équation différentielle) :

$$f'(x) = a(x)f(x) + \frac{x}{x^2-1} = \frac{-2x\ln|x|}{(x^2-1)^2} + \frac{x}{x^2-1} = \frac{x(x^2-1) - 2x\ln|x|}{(x^2-1)^2}\cdot$$

– *On a facilement $\lim\limits_{x\to 0} f'(x) = 0$ donc d'après le théorème de prolongement de la dérivée, f est dérivable en 0 et $f'(0) = 0$.*

– *Pour calculer $\lim\limits_{x\to 1} f'(x)$ on fait un rapide développement limité, en posant $x = 1+h$ et après avoir « simplifié » les termes qui ne posent pas de problème :*

$$f'(x) = \frac{x\left(x^2-1-2\ln|x|\right)}{(x+1)^2(x-1)^2} \underset{x\to 1}{\sim} \frac{x^2-1-2\ln x}{4(x-1)^2}$$

$$\lim_{x\to 1} f'(x) = \lim_{h\to 0} \frac{(1+h)^2 - 1 - 2\ln(1+h)}{4h^2}$$

$$= \lim_{h\to 0} \frac{h^2 + 2h - 2\left(h - \frac{h^2}{2} + O\left(h^3\right)\right)}{4h^2} = \lim_{h\to 0} \frac{2h^2 + O\left(h^3\right)}{4h^2} = \frac{1}{2}$$

donc d'après le théorème de prolongement de la dérivée, f est dérivable en 1 et $f'(1) = \dfrac{1}{2}$.

– *f étant impaire, f' est paire et on a le même résultat en -1.*

En conclusion, la seule solution de (E) sur $\mathbb{R}$ est l'application $x \mapsto \dfrac{x^2 \ln|x|}{x^2-1}$ (prolongée par continuité en -1, 0 et 1).

SF17.2 Reconnaître certaines situations particulières

Dans certains cas, un changement de variable ou de fonction inconnue permet de simplifier l'équation différentielle.
Dans d'autres cas, on peut « voir » une solution particulière sans avoir besoin d'utiliser la méthode de variation de la constante.

Exemple

Résoudre l'équation différentielle $2x\,\mathrm{e}^y\,y' + \mathrm{e}^y - x^2 = 0$.

On pense d'emblée au changement de fonction inconnue $z = \mathrm{e}^y$.
z est alors solution de l'équation différentielle :

$$(E) \quad 2xz' + z = x^2\,.$$

On résout alors (E) sur $I_1 =]-\infty\,;0[$ ou $I_2 =]0\,;+\infty[$.
Les solutions de l'équation homogène sont les fonctions :

$$z\colon x \mapsto \frac{\lambda_k}{\sqrt{|x|}}\,, \quad k \in \{1,2\}\,.$$

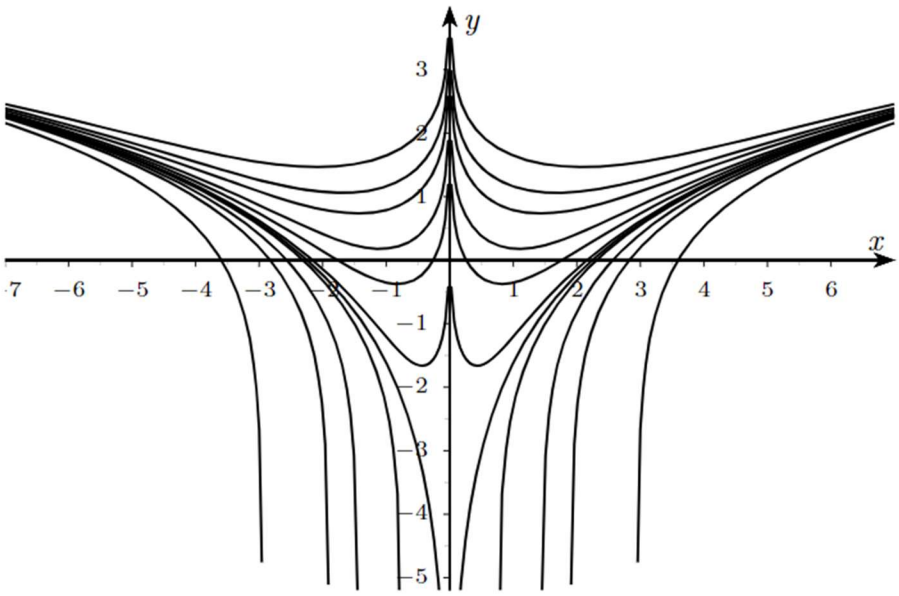
On pourrait utiliser la méthode de variation de la constante, mais un peu d'intuition fait penser à chercher une solution particulière polynomiale, forcément de degré 2.
On trouve facilement la fonction $x \mapsto \dfrac{x^2}{5}$, donc l'ensemble des solutions de (E) sur I_k est l'ensemble des fonctions de la forme :

$$z\colon x \mapsto \frac{\lambda_k}{\sqrt{|x|}} + \frac{x^2}{5}\,, \quad k \in \{1,2\}\,.$$

On en déduit que l'ensemble des solutions de l'équation initiale est l'ensemble des applications de la forme :

$$x \mapsto \ln\left(\frac{\lambda}{\sqrt{|x|}} + \frac{x^2}{5}\right)$$

ces fonctions n'étant définies que sur une partie de $\mathbb{R}$ (plus précisément, $\mathbb{R}^*$ si $\lambda \geqslant 0$, et $\left]-\infty\,;-(-5\lambda)^{2/5}\right[\cup \left](-5\lambda)^{2/5}\,;+\infty\right[$ si $\lambda < 0$).
Nous avons représenté certaines courbes intégrales ci-après.



■ Comment résoudre une équation différentielle scalaire linéaire d'ordre 2 ?

Nous ne traiterons pas ici des équations à coefficients constants, que l'on sait résoudre et qui ont été exposées en **[S17.8]**. Elles ont d'ailleurs été étudiées en 1re année.

SF17.3 Utiliser deux solutions de l'équation homogène

Le cas où l'on connaît deux solutions linéairement indépendantes y_1 et y_2 de l'équation homogène est évidemment le plus simple (même si le nombre de calculs peut être important), puisque l'on peut alors utiliser la méthode de variation des constantes.

Celle-ci consiste à chercher une solution de l'équation avec second membre sous la forme $y(x) = a(x)y_1(x) + b(x)y_2(x)$ où ici a et b sont des fonctions inconnues. Mais si vous contentez, comme on le voit très souvent, de remplacer y par cette expression dans l'équation, vous ne pourrez aboutir puisque vous aurez seulement *une* équation pour *deux* fonctions inconnues !

La méthode de variations des constantes, contrairement à celle de variation de *la* constante, ne peut donc pas s'improviser : il faut connaître les formules du cours ! Et pour retenir ces formules, il est important de se souvenir comment elles ont été obtenues (à partir du système différentiel associé à l'équation), ce qui permet, avec un peu de réflexion, de les retrouver rapidement.

Exemple

Soit l'équation différentielle : $t^2x'' - 2tx' + 2x = 2t^2$.

Chercher des solutions de l'équation homogène sur $\mathbb{R}^*$ sous la forme $t \mapsto t^n$ avec $n \in \mathbb{N}$.

En déduire l'ensemble des solutions de l'équation sur $\mathbb{R}^*$ puis sur $\mathbb{R}$.

- *Par un calcul simple, on vérifie que la fonction $t \mapsto t^n$ est solution sur $\mathbb{R}^*$ de l'équation homogène $t^2x'' - 2tx' + 2x = 0$ si et seulement si $n \in \{1,2\}$.*
- *Sur $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$, l'équation s'écrit sous forme réduite :*

$$x'' - \frac{2}{t}x' + \frac{2}{t^2} = 2\,.$$

Par la méthode de variations des constantes, on en cherche les solutions sous la forme :

$$x(t) = c_1(t)t + c_2(t)t^2\,,$$

où c_1, c_2 sont des fonctions de classe $\mathscr{C}^1$ telles que, pour $t \in \mathbb{R}_-^*$ ou $t \in \mathbb{R}_+^*$: $c_1'(t)\begin{pmatrix} t \\ 1 \end{pmatrix} + c_2'(t)\begin{pmatrix} t^2 \\ 2t \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$.

✓ *Plutôt que de retenir les formules du cours sous forme d'un système, il nous semble plus facile de les retenir en revenant à la méthode de variations des constantes pour le système différentiel linéaire associé, d'où la formulation ci-dessus.*

Cela conduit au système : $\begin{cases} tc_1'(t) + t^2c_2'(t) = 0 \\ c_1'(t) + 2tc_2'(t) = 2 \end{cases}$, *qui donne immédiatement :* $c_1'(t) = -2$ *et* $c_2'(t) = \dfrac{2}{t}$.
On obtient donc :

$$c_1(t) = -2t + a \quad \text{et} \quad c_2(t) = 2\ln|t| + b \quad \text{avec } (a,b) \in \mathbb{R}^2.$$

Finalement la solution générale de l'équation sur $\mathbb{R}^*$ est :

$$\begin{cases} x(t) = 2t^2\,(\ln(-t) - 1) + a_1t + a_2t^2 & \text{si } t < 0 \\ x(t) = 2t^2\,(\ln t - 1) + b_1t + b_2t^2 & \text{si } t > 0 \end{cases}$$

où a_1, a_2,b_1, b_2 sont des constantes réelles.

- *Toutes les solutions sont prolongeables par continuité en 0, avec une dérivée à gauche en 0 qui vaut a_1 et une dérivée à droite en 0 qui vaut b_1. La fonction prolongée est donc dérivable sur $\mathbb{R}$ si $a_1 = b_1$, mais $\lim\limits_{x\to 0} \dfrac{x'(t) - x'(0)}{t} = -\infty$: elle n'est pas deux fois dérivable en 0. Il n'y a donc pas de solutions sur $\mathbb{R}$.*

SF17.4 Utiliser une solution de l'équation homogène

Le cas où l'on connaît une solution de l'équation homogène est assez simple, puisque l'on peut alors utiliser la méthode de variation de la constante qui permet de se ramener à des équations différentielles linéaires du 1er ordre.
Contrairement à la méthode de variation *des* constantes, cette méthode est facile à retenir puisque le point de départ est le même que pour les équations du 1er ordre.

Exemple

Résoudre l'équation différentielle :

$$(E) \quad tx'' + 2(t+1)x' + 2x = 1 - \mathrm{e}^{-2t}\,,$$

en cherchant une solution de l'équation homogène sous la forme $t \mapsto t^\alpha$.

- *À l'aide de l'indication fournie, on trouve* $x = \dfrac{1}{t}$ *comme solution particulière de l'équation homogène (nous ne détaillons pas les calculs).*
- *Sur chaque intervalle* $\mathbb{R}_-^*$ *et* $\mathbb{R}_+^*$*, on cherche alors les solutions de l'équation sous la forme :* $x(t) = \dfrac{y(t)}{t}$*. Un calcul rapide donne :*

$$x \text{ solution de } (E) \iff y \text{ solution de } (E') : \quad y'' + 2y' = 1 - \mathrm{e}^{-2t}.$$

(E') est une équation différentielle linéaire du 1er ordre en y', à coefficients constants, et dont le second membre est d'une forme classique. Le cours permet d'en donner l'ensemble des solutions :

$$y' = \frac{1}{2} - t\mathrm{e}^{-2t} + \lambda \mathrm{e}^{-2t} \quad \text{avec } \lambda \in \mathbb{R}\,.$$

En intégrant on trouve : $y = \dfrac{t}{2} + \dfrac{t}{2}\mathrm{e}^{-2t} + C\mathrm{e}^{-2t} + D$ *avec* $C,\ D \in \mathbb{R}$*.*
Finalement les solutions de l'équation initiale sur $\mathbb{R}^*$ *sont de la forme :*

$$x(t) = \begin{cases} C_1 \dfrac{\mathrm{e}^{-2t}}{t} + \dfrac{D_1}{t} + \dfrac{1}{2}(1 + \mathrm{e}^{-2t}) & \text{si } t > 0 \\ C_2 \dfrac{\mathrm{e}^{-2t}}{t} + \dfrac{D_2}{t} + \dfrac{1}{2}(1 + \mathrm{e}^{-2t}) & \text{si } t < 0\,. \end{cases}$$

- *Pour que cette fonction soit prolongeable par continuité à droite en* 0*, il faut et il suffit que* $C_1 = -D_1$ *et pour qu'elle soit prolongeable par continuité à gauche en* 0*, il faut et il suffit que* $C_2 = -D_2$*.*
Dans ce cas, x s'écrit :

$$x(t) = \begin{cases} C_1 \left(\dfrac{\mathrm{e}^{-2t} - 1}{t}\right) + \dfrac{1}{2}(1 + \mathrm{e}^{-2t}) & \text{si } t > 0 \\ C_2 \left(\dfrac{\mathrm{e}^{-2t} - 1}{t}\right) + \dfrac{1}{2}(1 + \mathrm{e}^{-2t}) & \text{si } t < 0\,. \end{cases}$$

Il faut alors $C_1 = C_2$ *pour que cette fonction soit continue en* 0*. On a alors :*

$$\forall t \in \mathbb{R}^*\ ,\ x(t) = C\left(\frac{\mathrm{e}^{-2t} - 1}{t}\right) + \frac{1}{2}(1 + \mathrm{e}^{-2t}) \quad (C \in \mathbb{R})\,,$$

et ces fonctions prolongées par continuité en 0 *sont toutes solutions sur* $\mathbb{R}$ *car elles sont de classe* $\mathscr{C}^\infty$ *sur* $\mathbb{R}$*. Pour le justifier, il suffit de remarquer que la fonction* $t \mapsto \dfrac{\mathrm{e}^{-2t} - 1}{t}$ *est développable en série entière, de rayon de convergence* $+\infty$*.*

SF17.5 Utiliser un développement en série entière

On peut chercher en général une solution développable en série entière lorsque les coefficients sont des fonctions polynomiales.
On procède dans ce cas par analyse-synthèse : si une solution est développable en série entière, en injectant l'expression dans l'équation, on peut arriver à trouver les coefficients. Mais il ne faut pas oublier ensuite de vérifier que la solution obtenue convient, c'est-à-dire que le rayon de convergence de la série entière obtenue n'est pas nul.

Exemple

L'équation : $(E) : x^3y''+x(2x-1)y'+y = 1$ admet-elle des solutions développables en série entière ?

Si y est une telle solution, il existe $R > 0$ tel que :

$$\forall\, x \in]-R\,;R[,\ y(x) = \sum_{n=0}^{+\infty} a_n x^n\,.$$

Le théorème de dérivation d'une série entière permet d'écrire :

$$\forall\, x \in]-R\,;R[\ ,\ y'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} \quad \text{et} \quad y''(x) = \sum_{n=2}^{+\infty} n(n-1) a_n x^{n-2}\,.$$

y est solution de (E) sur $]-R\,;R[$ si et seulement si :

$$\sum_{\substack{n=0\\ \cancel{n=2}}}^{+\infty} n(n-1)a_n x^{n+1} + 2\sum_{\substack{n=0\\ \cancel{n=1}}}^{+\infty} n a_n x^{n+1} - \sum_{n=1}^{+\infty} n a_n x^n + \sum_{n=0}^{+\infty} a_n x^n = 1\,,$$

soit :

$$\sum_{n=0}^{+\infty} n(n+1)a_n x^{n+1} - \sum_{n=1}^{+\infty} (n-1)a_n x^n + a_0 = 1\,,$$

ou, après changement d'indices :

$$\sum_{n=0}^{+\infty} n(n+1)a_n x^{n+1} - \sum_{n=0}^{+\infty} n a_{n+1} x^{n+1} + a_0 = 1\,.$$

Par unicité du développement en série entière, on obtient :

$$a_0 = 1 \quad \text{et} \quad \forall\, n \in \mathbb{N},\ n(n+1)a_n = n a_{n+1}$$

c'est-à-dire :

$$a_0 = 1 \quad \text{et} \quad \forall\, n \in \mathbb{N}^*,\ a_{n+1} = (n+1)a_n\,.$$

Il est facile d'en déduire par récurrence : $\forall\, n \in \mathbb{N}^*,\ a_n = n!a_1$.

Mais la série entière $\sum_{n\geqslant 1} n!x^n$ a un rayon de convergence nul (car la suite $(n!x^n)$ n'est pas bornée pour $x \neq 0$), donc ne peut être solution. On doit donc avoir $a_1 = 0$. Dans ce cas, tous les a_n pour $n \geqslant 1$ sont nuls, et il ne reste que $a_0 = 1$.

En conclusion, l'équation admet pour seule solution développable en série entière la fonction constante égale à 1 (qui était dès le début une solution évidente).

Exemple

Résoudre l'équation différentielle : $(E) : 4xy'' + 2y' - y = 0$.

• *Supposons qu'il existe une série entière de rayon de convergence $R > 0$ solution de (E) sur $\mathbb{R}$. Notons : $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.*

Pour les mêmes raisons que dans l'exemple précédent, f est solution de (E) sur $]-R\,;R[$ si et seulement si :

$$4x\sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2} + 2\sum_{n=1}^{+\infty} n a_n x^{n-1} - \sum_{n=0}^{+\infty} a_n x^n = 0\,,$$

d'où, après changement d'indices :

$$\sum_{\cancel{n=1}\ n=0}^{+\infty} 4(n+1)n a_{n+1} x^n + \sum_{n=0}^{+\infty} 2(n+1)a_{n+1}x^n - \sum_{n=0}^{+\infty} a_n x^n = 0$$

(on peut faire commencer la première somme à $n = 0$ puisque le terme correspondant est nul). On a donc :

$$\sum_{n=0}^{+\infty} \big((2n+2)(2n+1)a_{n+1} - a_n\big)x^n = 0\,.$$

Par unicité du développement en série entière :

$$\forall\, n \in \mathbb{N},\ a_{n+1} = \frac{a_n}{(2n+1)(2n+2)}\,\cdot$$

On en déduit par une récurrence facile :

$$\forall\, n \in \mathbb{N},\ a_n = \frac{a_0}{(2n)!}\,\cdot$$

Ainsi, si (E) admet des solutions développables en séries entières, elles sont toutes proportionnelles à la fonction f définie par :

$$\forall\, x \in\,]-R\,;R[\ ,\ f(x) = \sum_{n=0}^{+\infty} \frac{x^n}{(2n)!}\,\cdot$$

Puisque, pour tout $x \in \mathbb{R}$, $\lim\limits_{n\to+\infty} \dfrac{x^n}{(2n)!} = 0$, on a $R = +\infty$. La fonction f définie ci-dessus est donc de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$, et elle est forcément solution de (E), d'après les calculs faits.

De plus, on sait que :

$$\forall\, x \in \mathbb{R},\ \operatorname{ch}\ x = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!} \quad \textit{et} \quad \cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

donc

$$\forall\, x \in \mathbb{R}_+\ ,\ \operatorname{ch}\ \sqrt{x} = \sum_{n=0}^{\infty} \frac{x^n}{(2n)!} \quad \textit{et} \quad \forall\, x \in \mathbb{R}_-\ ,\ \cos\sqrt{-x} = \sum_{n=0}^{\infty} \frac{x^n}{(2n)!}\,\cdot$$

Ainsi : $f(x) = \begin{cases} \operatorname{ch}\sqrt{x} & \textit{si } x \geqslant 0 \\ \cos\sqrt{-x} & \textit{si } x < 0\,. \end{cases}$

• Résolution de (E) sur $]0\,;+\infty[$

Soit y une solution quelconque de (E) sur $]0\,;+\infty[$. Posons $y = zf$ (ce qui est possible car f ne s'annule pas sur $\mathbb{R}_+$), et recherchons une équation différentielle (E') vérifiée par z (méthode de variation de la constante). On a :

$$y' = z'f + zf' \quad \textit{et} \quad y'' = z''f + 2z'f' + zf''$$

puis en remplaçant dans (E), on obtient :

$$4x(z''f + 2z'f' + zf'') + 2(z'f + zf') - zf = 0\,,$$

et comme f est solution de (E) : $z(4xf'' + 2f' - f) = 0$ donc il reste finalement :

$$4xf(x)z'' + (8xf'(x) + 2f(x))z' = 0 \quad \textit{soit} \quad z'' + \left(2\frac{f'(x)}{f(x)} + \frac{1}{2x}\right)z' = 0\,.$$

Comme sur $\mathbb{R}_+^*$ on a : $f(x) = \operatorname{ch}\sqrt{x}$, on obtient l'équation différentielle du premier ordre en z' :

$$(E') : z'' + \left(\frac{\operatorname{th}\sqrt{x}}{\sqrt{x}} + \frac{1}{2x}\right)z' = 0\,.$$

Une primitive de $x \mapsto \dfrac{\operatorname{th}\sqrt{x}}{\sqrt{x}}$ est $x \mapsto 2\ln(\operatorname{ch}\sqrt{x})$, d'où :

$$z'(x) = A\mathrm{e}^{-2\ln(\operatorname{ch}\sqrt{x}) - \frac{1}{2}\ln x} = \frac{A}{(\operatorname{ch}\sqrt{x})^2\sqrt{x}} \quad (A \in \mathbb{R}).$$

Une primitive de $x \mapsto \dfrac{1}{(\operatorname{ch}\sqrt{x})^2\sqrt{x}}$ est $x \mapsto 2\operatorname{th}\sqrt{x}$, donc :

$$z(x) = 2A\operatorname{th}\sqrt{x} + B \quad (B \in \mathbb{R}).$$

Enfin : $y(x) = z(x)f(x) = (2A\ \operatorname{th}\sqrt{x} + B)\operatorname{ch}\sqrt{x} = 2A\operatorname{sh}\sqrt{x} + B\operatorname{ch}\sqrt{x}$.
Les solutions de (E) sur $\mathbb{R}_+^*$ sont donc de la forme :

$$y(x) = A\operatorname{ch}\sqrt{x} + B\operatorname{sh}\sqrt{x} \text{ avec } A,\ B \in \mathbb{R}\,.$$

• Résolution de (E) sur $]-\infty\,;0[$

On pourrait procéder de la même façon sur $\mathbb{R}_-^*$. Mais, comme on se doute un peu de la forme des solutions, il est plus rapide de vérifier directement par le calcul que la fonction $x \mapsto \sin\sqrt{-x}$ est solution de l'équation. Puisque l'on sait que l'ensemble des solutions sur $\mathbb{R}_-^*$ forme un espace vectoriel de dimension 2, on peut dire directement qu'il s'agit de l'ensemble des applications de la forme :

$$y(x) = C\cos\sqrt{-x} + D\sin\sqrt{-x} \text{ avec } C,\ D \in \mathbb{R}\,.$$

• Recherche d'une solution sur $\mathbb{R}$ (raccordement)

Soit y une solution sur $\mathbb{R}$. Elle est donc de la forme

$$y(x) = \begin{cases} A\operatorname{ch}\sqrt{x} + B\operatorname{sh}\sqrt{x} & \text{si } x > 0 \\ C\cos\sqrt{-x} + D\sin\sqrt{-x} & \text{si } x < 0\,. \end{cases}$$

Nous vous laissons le soin de vérifier que le raccordement par continuité en 0 exige $A = C$, et que la dérivabilité en 0 exige ensuite $B = D = 0$.
Cela montre que les seules solutions définies sur $\mathbb{R}$ sont exactement celles développables en série entière trouvées au début.

SF17.6 Utiliser un changement de variable

La technique du changement de variable dans une équation différentielle est un peu délicate, et nous vous conseillons de bien étudier les calculs dans l'exemple qui suit, et de les refaire.

Exemple

Résoudre l'équation différentielle : $x^2y'' + 4xy' + 2y = 0 \quad (H)$, en posant $x = \pm\mathrm{e}^t$.

Le coefficient x^2 de y'' s'annulant en 0, on est amené à résoudre l'équation séparément sur $\mathbb{R}_-^*$ et sur $\mathbb{R}_+^*$.

- Étude sur $\mathbb{R}_+^*$

  Comme l'énoncé le suggère, on considère ici le changement de variable $x = \mathrm{e}^t$.
  L'application $\varphi\colon t \mapsto \mathrm{e}^t$ étant une bijection de classe $\mathscr{C}^2$ de $\mathbb{R}$ sur $\mathbb{R}_+^*$, on pourra poser $y(x) = y\circ\varphi(t) = z(t)$, et y est de classe $\mathscr{C}^2$ si et seulement si il en est de même de z. Il reste à trouver la nouvelle équation différentielle vérifiée par z. On calcule :

$$z'(t) = \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\big(y(x)\big) = \frac{\mathrm{d}y}{\mathrm{d}x}.\frac{\mathrm{d}x}{\mathrm{d}t} = y'(x).\mathrm{e}^t\,,$$

  puis

$$\begin{aligned} z''(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\big(y'(x).\mathrm{e}^t\big) = \frac{\mathrm{d}}{\mathrm{d}t}\big(y'(x)\big)\mathrm{e}^t + y'(x).\mathrm{e}^t \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\big(y'(x)\big)\frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{e}^t + y'(x).\mathrm{e}^t = y''(x)\big(\mathrm{e}^t\big)^2 + y'(x)\mathrm{e}^t\,, \end{aligned}$$

  soit, en abrégé : $z' = xy'$ et $z'' = x^2y'' + xy'$.

  ✓ *Nous avons ici utilisé les notations différentielles pour faire les calculs, à la manière des physiciens. Une autre écriture, peut-être plus rigoureuse, est la suivante :*

$$z'(t) = (y\circ\varphi)'(t) = \varphi'(t)\cdot(y'\circ\varphi)(t) = \mathrm{e}^t\cdot y'(x) = xy'(x)$$

puis

$$z''(t) = \big(\varphi'\cdot(y'\circ\varphi)\big)'(t) = \varphi''(t)\cdot(y'\circ\varphi)(t) + \big(\varphi'(t)\big)^2\cdot(y''\circ\varphi)(t)$$

soit

$$z''(t) = \mathrm{e}^t\cdot y'(x) + (\mathrm{e}^t)^2\cdot y''(x) = xy'(x) + x^2y''(x)\,.$$

  À vous de choisir l'écriture avec laquelle vous vous sentez le plus à l'aise.

  Finalement, y vérifie l'équation (H) si et seulement si z vérifie l'équation :

$$z'' + 3z' + 2z = 0 \quad (H')$$

  Il s'agit là d'une équation différentielle linéaire à coefficients constants. Son équation caractéristique est : $r^2 + 3r + 2 = 0$, de racines distinctes $r = -1$ et $r = -2$. L'ensemble des solutions est donc l'ensemble des fonctions z de la forme :

$$z(t) = a\mathrm{e}^{-t} + b\mathrm{e}^{-2t} \quad \textit{avec } (a,b)\in\mathbb{R}^2\,,$$

  et, puisque $x = \mathrm{e}^t$, l'ensemble des solutions sur $\mathbb{R}_+^*$ de l'équation initiale est l'ensemble des applications y de la forme :

$$y(x) = \frac{a}{x} + \frac{b}{x^2} \quad \textit{avec } (a,b)\in\mathbb{R}^2\,.$$

- Étude sur $\mathbb{R}_-^*$

  On pourrait faire de même sur $\mathbb{R}_-^*$, en considérant cette fois-ci le changement de variable $x = -\mathrm{e}^t$.

Mais il est plus astucieux de remarquer que, comme le calcul des dérivées se fait ici de la même façon sur $\mathbb{R}_+^*$ et sur $\mathbb{R}_-^*$, les applications $x \mapsto \frac{a}{x} + \frac{b}{x^2}$ sont encore solutions sur $\mathbb{R}_-^*$. Puisque l'ensemble des solutions sur $\mathbb{R}_-^*$ est, d'après le théorème de Cauchy-Lipschitz, un $\mathbb{R}$-espace vectoriel de dimension 2, on aura bien ainsi l'ensemble des solutions.

En conclusion, l'ensemble des solutions de l'équation (H) est l'ensemble des applications y de la forme :

$$y(x) = \begin{cases} \dfrac{a}{x} + \dfrac{b}{x^2} & \text{si } x < 0 \\ \dfrac{c}{x} + \dfrac{d}{x^2} & \text{si } x > 0 \end{cases} \qquad \text{avec } (a,b,c,d) \in \mathbb{R}^4 .$$

SF17.7 Utiliser un changement de fonction inconnue

Il ne faut pas confondre cette méthode avec la précédente, et elle est beaucoup plus facile à comprendre et à mettre en œuvre. C'est d'ailleurs cette méthode que l'on utilise lorsqu'on fait la méthode de variation de la constante.

Exemple

Résoudre l'équation différentielle :

$$xy'' - 2(x-1)y' + (x-2)y = x\mathrm{e}^x \quad (E)$$

à l'aide du changement de fonction inconnue $z = y' - y$.

- *Notons déjà, comme d'habitude, que l'on se place pour ce qui suit sur l'un des intervalles $\mathbb{R}_-^*$ ou $\mathbb{R}_+^*$, où le coefficient de y'' ne s'annule pas.*
- *Avec la notation de l'énoncé, si y est solution de (E), alors z est de classe $\mathscr{C}^1$ et vérifie l'équation différentielle :*

$$xz' - (x-2)z = x\mathrm{e}^x \quad (E')$$

La solution générale de l'équation homogène associée :

$$xz' = (x-2)z \quad (H)$$

est : $x \mapsto C\exp\left(\displaystyle\int \frac{x-2}{x}\,\mathrm{d}x\right) = C\exp\left(x - 2\ln|x|\right) = C\frac{\mathrm{e}^x}{x^2}$.

On cherche alors la solution générale de (E') par la méthode de variation de la constante, en posant $z(x) = C(x)\frac{\mathrm{e}^x}{x^2}$ (il n'y a pas de problème puisque la fonction $x \mapsto \frac{\mathrm{e}^x}{x^2}$ ne s'annule pas), ce qui conduit à :

$$xC'(x)\frac{\mathrm{e}^x}{x^2} = x\mathrm{e}^x \quad \text{soit} \quad C'(x) = x^2 .$$

La solution générale de (E') est donc :

$$z(x) = \frac{x\mathrm{e}^x}{3} + \lambda\frac{\mathrm{e}^x}{x^2} .$$

- *Il ne reste plus qu'a résoudre l'équation différentielle :*

$$y' - y = \frac{x\mathrm{e}^x}{3} + \lambda\frac{\mathrm{e}^x}{x^2} .$$

Une solution de l'équation homogène associée étant $x \mapsto \mathrm{e}^x$, on cherche donc la solution sous la forme $y(x) = \varphi(x)\mathrm{e}^x$ ce qui conduit à

$$\varphi'(x)\mathrm{e}^x = \frac{x\mathrm{e}^x}{3} + \lambda\frac{\mathrm{e}^x}{x^2} \quad \textit{soit} \quad \varphi'(x) = \frac{x}{3} + \frac{\lambda}{x^2}$$

puis $\varphi(x) = \dfrac{x^2}{6} - \dfrac{\lambda}{x} + \mu$ *et enfin* $y(x) = \dfrac{x^2}{6}\mathrm{e}^x - \dfrac{\lambda}{x}\mathrm{e}^x + \mu\mathrm{e}^x$.

• *Les solutions de l'équation (E) sur $\mathbb{R}^*$ sont finalement de la forme :*

$$y(x) = \begin{cases} \dfrac{x^2}{6}\mathrm{e}^x - \dfrac{\lambda_1}{x}\mathrm{e}^x + \mu_1\mathrm{e}^x & \textit{si } x < 0 \\ \dfrac{x^2}{6}\mathrm{e}^x - \dfrac{\lambda_2}{x}\mathrm{e}^x + \mu_2\mathrm{e}^x & \textit{si } x > 0 \end{cases} \qquad \textit{avec } (\lambda_1,\mu_1,\lambda_2,\mu_2) \in \mathbb{R}^4 .$$

Pour qu'une telle solution se prolonge par continuité en 0, il faut et il suffit que $\lambda_1 = \lambda_2 = 0$ et $\mu_1 = \mu_2$. Les solutions définies sur $\mathbb{R}$ sont donc les applications de la forme $x \mapsto \dfrac{x^2}{6}\mathrm{e}^x + \mu\mathrm{e}^x$ (elles sont bien de classe $\mathscr{C}^\infty$).

■ Comment résoudre un système différentiel ?

Sauf quelques rares cas, on ne sait résoudre un système différentiel que lorsqu'il est à coefficients constants. La seule méthode générale consiste alors à réduire la matrice du système, c'est-à-dire à la diagonaliser ou à la trigonaliser.

SF17.8 Diagonaliser ou trigonaliser la matrice

Si on a un système différentiel linéaire homogène de la forme $X' = AX$, où A est diagonalisable, on a vu en **[S17.5]** que l'on sait en exprimer les solutions dès que l'on connaît les valeurs propres de A et une base de vecteurs propres associés.

Mais il nous semble inutile de retenir cette formule (il y en a d'autres bien plus importantes !), car le principe de résolution est simple : si $B = P^{-1}AP$, le système $X' = AX$ équivaut à $P^{-1}X' = BP^{-1}X$, donc, en posant $Y(t) = P^{-1}X(t)$, soit $X(t) = PY(t)$ (le calcul de P^{-1} n'est donc pas utile en pratique), on obtiendra le nouveau système différentiel $Y' = BY$ (en effet, $X(t) = PY(t)$ implique $X'(t) = PY'(t)$ car la matrice P est constante).
De plus, dans le cas où la matrice de passage est constante, cette méthode permet aussi de résoudre certains systèmes à coefficients non constants.

Exemple

Résoudre le système différentiel : $\begin{cases} x'(t) = (t-2)x(t) - (t-1)y(t) \\ y'(t) = 2(t-1)x(t) - (2t-1)y(t) . \end{cases}$

En posant comme d'habitude $X = \begin{pmatrix} x \\ y \end{pmatrix}$, ce système s'écrit $X' = A(t)X$ avec : $A(t) = \begin{pmatrix} t-2 & -t+1 \\ 2t-2 & -2t+1 \end{pmatrix}$

Nous ne détaillons pas ici les calculs ; on trouve assez facilement que la matrice $A(t)$ est diagonalisable, et que $A(t) = PD(t)P^{-1}$ avec :

$$D(t) = \begin{pmatrix} -t & 0 \\ 0 & -1 \end{pmatrix} \quad \textit{et} \quad P = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

La matrice P ne dépendant pas de t, on peut appliquer la méthode exposée plus haut.
Posons $Y(t) = P^{-1}X(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$, soit $X(t) = PY(t)$. $Y(t)$ est alors solution du système $Y' = D(t)Y$ soit :

$$\begin{cases} u'(t) = -t\,u(t) \\ v'(t) = -v(t) \end{cases}$$

D'où $u(t) = C_1 \mathrm{e}^{-\frac{t^2}{2}}$ et $v(t) = C_2 \mathrm{e}^{-t}$, avec C_1, C_2 constantes.
Puisque $X(t) = PY(t)$, on trouve finalement :

$$x(t) = C_1 \mathrm{e}^{-\frac{t^2}{2}} + C_2 \mathrm{e}^{-t} \quad \text{et} \quad y(t) = 2C_1 \mathrm{e}^{-\frac{t^2}{2}} + C_2 \mathrm{e}^{-t}\,.$$

Exemple

Résoudre le système différentiel : $\begin{cases} x' = 2x - y + 2z \\ y' = 10x - 5y + 7z \\ z' = 4x - 2y + 2z\,. \end{cases}$

C'est un système différentiel linéaire d'ordre 1 homogène qui s'écrit matriciellement $X' = AX$ avec : $A = \begin{pmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{pmatrix}$.
On en calcule le polynôme caractéristique : $\chi_A(X) = -X^2(X+1)$.
A n'est pas diagonalisable ; il faut donc la trigonaliser et après calculs on peut trouver $A = PTP^{-1}$ avec :

$$P = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 1 \\ 2 & 0 & 1 \end{pmatrix} \quad \text{et} \quad T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\,.$$

✓ *Nous n'avons pas détaillé les calculs ; pour la façon de trigonaliser une matrice, reportez-vous à* **[SF4.15]**. *Cet exemple vous permettra de vous exercer.*

En posant $Y = P^{-1}X$, le système équivaut à $Y' = TY$ soit à

$$\begin{cases} y_1' = -y_1 \\ y_2' = \ y_3 \\ y_3' = \ 0 \end{cases}$$

qui donne facilement $Y = \begin{pmatrix} \lambda \mathrm{e}^{-t} \\ \mu t + \nu \\ \mu \end{pmatrix}$ avec λ, μ, $\nu \in \mathbb{K}$.
La solution générale du système est donnée par $X = PY$ soit :

$$X(t) = \lambda \begin{pmatrix} -\mathrm{e}^{-t} \\ \mathrm{e}^{-t} \\ 2\mathrm{e}^{-t} \end{pmatrix} + \mu \begin{pmatrix} t \\ 2t+1 \\ 1 \end{pmatrix} + \nu \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \text{avec } \lambda,\ \mu,\ \nu \in \mathbb{K}\,.$$

✓ Dans ces deux exemples, vous remarquerez, comme nous vous l'avions annoncé, qu'il est inutile de calculer P^{-1} dans le cas d'un système homogène (c'est en revanche nécessaire si il y a un second membre).

SF17.9 Utiliser des combinaisons linéaires

Parfois, certaines combinaisons linéaires astucieuses peuvent permettre de simplifier le système.

Exemple

Résoudre le système différentiel : $\begin{cases} x' = \frac{1}{t}y + 1 \\ y' = \frac{1}{t}x + 1 \end{cases}$, où x et y sont deux fonctions dérivables sur $\mathbb{R}_+^*$.

En posant : $u = x + y$ et $v = x - y$, on obtient le nouveau système :

$$\begin{cases} u' = \frac{1}{t}u + 2 \\ v' = -\frac{1}{t}v\,. \end{cases}$$

On est donc ramené à résoudre deux équations différentielles scalaires d'ordre 1. On obtient facilement :

$$\forall\, t \in \mathbb{R}_+^*,\ u(t) = (2\ln t + C_1)t \quad \text{et} \quad v(t) = \frac{C_2}{t} \quad \text{avec } (C_1,C_2) \in \mathbb{R}^2$$

puis finalement

$$\begin{cases} x(t) = \dfrac{u(t)+v(t)}{2} = t\ln t + \lambda t + \dfrac{\mu}{t} \\ y(t) = \dfrac{u(t)+v(t)}{2} = t\ln t + \lambda t - \dfrac{\mu}{t} \end{cases}$$

(en ayant posé $\lambda = \frac{C_1}{2}$ et $\mu = \frac{C_2}{2}$).

SF17.10 Utiliser l'exponentielle de matrice

Soit un système différentiel linéaire homogène de la forme $X' = AX$, où $A \in \mathcal{M}_n(\mathbb{K})$ est constante. D'après le cours, l'ensemble des solutions est l'ensemble :

$$\mathscr{S}_H = \left\{ t \mapsto \mathrm{e}^{tA}C \;\middle|\; C \in \mathcal{M}_{n,1}(\mathbb{K}) \right\}.$$

Mais sauf si le calcul de $\exp(A)$ est simple, on évitera d'utiliser cette formule dans la pratique. En effet, pour calculer $\exp(A)$, on est amené en général à diagonaliser ou trigonaliser A, et dans ce cas autant utiliser directement la méthode exposée en **[SF17.8]**.

Exemple

Résoudre le système différentiel : $X' = AX$ avec $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$.

A est de rang 1, et on vérifie facilement que $A^2 = 0_3$.
Donc $\exp(tA) = \mathrm{I}_3 + tA$, et l'ensemble des solutions est l'ensemble des fonctions $X : t \mapsto (\mathrm{I}_3 + tA)\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}$ avec $(\lambda,\mu,\nu) \in \mathbb{R}^3$.

L'exponentielle de matrice peut par contre être très utile dans certains problèmes théoriques.

Exemple

Soit l'équation différentielle : $(H)\quad X' = AX$, avec $A \in \mathcal{M}_n(\mathbb{C})$ constante.
Montrer que les propositions suivantes sont équivalentes :

(i) toute solution de (H) est bornée sur $\mathbb{R}$,

(ii) A est diagonalisable et $\operatorname{Sp} A \subset \mathrm{i}\mathbb{R}$.

- $(ii) \implies (i)$
  Supposons A diagonalisable de valeurs propres $\lambda_1, \ldots, \lambda_n$ imaginaires pures. Si $(V_1, \ldots, V_n)$ est une base de vecteurs propres associés, on a :
  $$\forall t \in \mathbb{R},\ X(t) = \sum_{k=1}^{n} \alpha_k \mathrm{e}^{\lambda_k t} V_k \quad \text{où } (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n.$$
  Si $\|\cdot\|$ est une norme quelconque sur $\mathcal{M}_{n,1}(\mathbb{C})$, on a par l'inégalité triangulaire :
  $$\forall t \in \mathbb{R},\ \|X(t)\| \leqslant \sum_{k=1}^{n} |\alpha_k| \left|\mathrm{e}^{\lambda_k t}\right| \|V_k\| ,$$
  et le résultat découle alors du fait que $\left|\mathrm{e}^{\lambda_k t}\right| = 1$ pour tout t puisque λ_k est imaginaire pur.
- $(i) \implies (ii)$
  Supposons que les solutions de (H) sont bornées.
  On sait d'après le cours que ces solutions sont de la forme :
  $$t \mapsto X(t) = \mathrm{e}^{tA} X_0 \quad \text{avec } X_0 \in \mathcal{M}_{n,1}(\mathbb{C}).$$
  - *Soit λ une valeur propre de A, et $V \in \mathcal{M}_{n,1}(\mathbb{C})$ un vecteur propre associé. Alors $AV = \lambda V$, d'où $A^k V = \lambda^k V$ pour tout $k \in \mathbb{N}$ puis $\mathrm{e}^{tA} V = \mathrm{e}^{\lambda t} V$.*
    Puisque $t \mapsto \mathrm{e}^{tA} V$ est solution de (H) elle est bornée, donc $t \mapsto \mathrm{e}^{\lambda t} V$ est bornée. Puisque $V \neq 0$, cela implique que la fonction $t \mapsto \mathrm{e}^{\lambda t}$ est bornée sur $\mathbb{R}$, donc λ est imaginaire pur (car si $\lambda = a + \mathrm{i}b$ avec a, b réels, $\left|\mathrm{e}^{\lambda t}\right| = \left|\mathrm{e}^{at}\mathrm{e}^{\mathrm{i}t}\right| = \mathrm{e}^{at}$).
  - *Pour montrer que A est diagonalisable, il suffit de démontrer que son polynôme minimal π_A (qui est scindé dans $\mathbb{C}[X]$) n'admet que des racines simples.*
    Raisonnons par l'absurde et supposons qu'il existe une valeur propre λ de A qui est racine d'ordre $m \geqslant 2$ de π_A. Alors il existe $V \in \mathcal{M}_{n,1}(\mathbb{C})$ tel que :
    $$V \in \operatorname{Ker}\big((A - \lambda \mathrm{I}_n)^m\big) \quad \text{et} \quad V \notin \operatorname{Ker}(A - \lambda \mathrm{I}_n)\,.$$
    Alors pour tout $t \in \mathbb{R}$,
    $$\mathrm{e}^{tA} V = \mathrm{e}^{\lambda t} \mathrm{e}^{t(A - \lambda \mathrm{I}_n)} V = \mathrm{e}^{\lambda t} \left(\sum_{k=0}^{+\infty} \frac{t^k}{k!} (A - \lambda \mathrm{I}_n)^k V \right)$$
    $$= \mathrm{e}^{\lambda t} V + \mathrm{e}^{\lambda t} \left(\sum_{k=1}^{m-1} \frac{t^k}{k!} (A - \lambda \mathrm{I}_n)^k V \right).$$

Or la fonction $t \mapsto \mathrm{e}^{tA}V$ est bornée puisque c'est une solution de (H), et la fonction $t \mapsto \mathrm{e}^{\lambda t}$ est aussi bornée puisque λ est imaginaire pur.
La fonction $t \mapsto \sum_{k=1}^{m-1} \frac{t^k}{k!}(A-\lambda \mathrm{I}_n)^k V$ doit donc aussi être bornée sur $\mathbb{R}$, ce qui est impossible puisque, $(A-\lambda \mathrm{I}_n)V$ n'étant pas nul, une au moins de ses composantes est une fonction polynôme en t de degré $\geqslant 1$.

Cela achève la démonstration de la seconde implication.

■ Comment obtenir des résultats théoriques sur les solutions d'une équation différentielle ?

Puisque l'on ne sait résoudre complètement que très peu d'équations différentielles, les problèmes les concernant peuvent être assez difficiles. Ils utilisent en général toutes les parties de l'analyse (intégration bien sûr, mais aussi séries de fonctions, séries entières, résultats sur la dérivation etc...).
Le seul résultat théorique qui concerne les équations différentielles est le théorème de Cauchy-Lipschitz, que l'on utilise donc très fréquemment.
Dans le cas d'équation du 1er ordre, il ne faut pas oublier que l'on connaît la forme générale des solutions.
On peut aussi dans certains cas s'inspirer de la méthode de variation de la (ou des) constante(s).

SF17.11 Exprimer les solutions

Lorsque l'on sait exprimer les solutions, même à l'aide d'une intégrale qu'on ne sait pas calculer, les problèmes sont grandement facilités. Nous avons d'ailleurs déjà utilisé cette méthode pour un système différentiel dans l'exemple précédent.

Exemple

Soit $a : \mathbb{R}_+ \to \mathbb{R}$ continue et intégrable sur $\mathbb{R}_+$.

Montrer que les solutions de l'équation différentielle $y' - a(x)y = 0$ sont bornées sur $\mathbb{R}_+$.

On sait d'après le cours que la solution générale de 1'équation considérée est donnée par :

$$y(x) = \lambda \mathrm{e}^{A(x)} \quad \text{avec } A(x) = \int_0^x a(t)\,\mathrm{d}t \quad \text{et} \quad \lambda \in \mathbb{R}\,.$$

Or, pour tout $x \geqslant 0$, $|A(x)| \leqslant \int_0^x |a(t)|udt \leqslant \int_0^{+\infty} |a(t)|\,\mathrm{d}t$, donc la fonction y est bornée.

SF17.12 Utiliser la méthode de variation des (ou de la) constante(s)

On peut s'inspirer de la méthode de variations des (ou de la) constantes(s) pour résoudre des problèmes qui « ressemblent » à des équations différentielles.

Exemple

Soit a une application continue définie sur un intervalle I à valeurs dans $\mathbb{R}$.
Montrer que si $f \in \mathscr{C}^1(I,\mathbb{R})$ vérifie :

$$\forall\, x \in I,\ f'(x) + a(x)f(x) \neq 0 \quad (*)$$

alors f s'annule au plus une fois sur I.

En notant A une primitive de la fonction a sur I, et en posant $f(x) = \varphi(x)\mathrm{e}^{-A(x)}$, la fonction $\varphi = f\mathrm{e}^{A}$ est de classe $\mathscr{C}^1$ sur I et, puisque $f' = \varphi'\mathrm{e}^{-A} - a\varphi$, l'hypothèse $(*)$ se traduit par :

$$\forall\, x \in I,\ \varphi'(x) \neq 0\,.$$

La fonction φ' étant continue, il résulte du théorème des valeurs intermédiaires qu'elle garde un signe constant sur I. Elle est donc strictement monotone sur I, donc s'annule au plus une fois.
Puisque $f = \mathrm{e}^{-A}\varphi$, il en est de même de f.

Exemple

Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$ vérifiant :

$$\forall\, x \in \mathbb{R},\ f''(x) + f(x) \geqslant 0\,.$$

Montrer que $f(x) + f(x+\pi) \geqslant 0$ pour tout $x \in \mathbb{R}$.

Posons $g(x) = f(x) + f''(x)$ pour tout $x \in \mathbb{R}$. Alors g est continue sur $\mathbb{R}$, à valeurs dans $\mathbb{R}_+$, et f est donc solution de l'équation différentielle :

$$y'' + y = g\,.$$

On utilise alors la méthode de variations des constantes : il existe u, v de classe $\mathscr{C}^1$ telles que :

$$\left\{\begin{array}{rcl} f(x) &=& u(x)\cos x + v(x)\sin x \\ f'(x) &=& -u(x)\sin x + v(x)\cos x \end{array}\right.$$

où u', v' vérifient le système :

$$\left\{\begin{array}{rcl} u'(x)\cos x + v'(x)\sin x &=& 0 \\ -u'(x)\sin x + v'(x)\cos x &=& g(x)\,. \end{array}\right.$$

On en déduit facilement $u'(x) = -g(x)\sin x$ et $v'(x) = g(x)\cos x$, donc il existe des constantes λ et μ telles que pour tout $x \in \mathbb{R}$:

$$u(x) = \lambda - \int_0^x g(t)\sin t\,\mathrm{d}t \quad \text{et} \quad v(x) = \mu + \int_0^x g(t)\cos t\,\mathrm{d}t\,.$$

On en déduit finalement :

$$\forall\, x \in \mathbb{R},\ f(x) = \lambda\cos x + \mu\sin x + \int_0^x g(t)\sin(x-t)\,\mathrm{d}t\,.$$

Par conséquent, pour tout x réel :

$$\begin{aligned} f(x) + f(x+\pi) &= \int_0^x g(t)\sin(x-t)\,\mathrm{d}t + \int_0^{x+\pi} g(t)\sin(x+\pi-t)\,\mathrm{d}t \\ &= \int_{x+\pi}^{x} g(t)\sin(x-t)\,\mathrm{d}t = \int_0^{\pi} g(x-u)\sin u\,\mathrm{d}u \geqslant 0 \end{aligned}$$

grâce à la positivité de g et celle de sin *sur $[0\,;\pi]$.*

SF17.13 Utiliser le théorème de Cauchy-Lipschitz

Les méthodes exposées ci-dessus utilisent le fait que, dans certains cas, on sait exprimer la forme générale des solutions d'une équation différentielle. Malheureusement, ces cas sont en nombre assez restreint ; le seul résultat qui reste alors à notre disposition est le théorème de Cauchy-Lipschitz (et ses conséquences).

Exemple

Soit p et q deux applications continues sur un intervalle I, à valeurs dans $\mathbb{R}$.
Soit f une solution sur I de l'équation différentielle $y'' + py' + qy = 0$, non nulle.

Montrer que si f s'annule en un point $a \in I$, il existe un voisinage de a dans I sur lequel f ne s'annule qu'en a (on dit que les zéros de f sont *isolés*).

Si $f(a) = 0$, on ne peut pas avoir $f'(a) = 0$ car sinon, d'après le théorème de Cauchy-Lipschitz, f serait la solution nulle.
Si par exemple $f'(a) > 0$, la continuité de f' implique qu'elle reste à valeurs strictement positives sur un voisinage de a. f est donc strictement croissante sur ce voisinage, donc ne peut s'y annuler qu'en a.

Exemple

Soit p et q deux applications continues de $\mathbb{R}$ dans $\mathbb{R}$. On suppose p impaire et q paire, et on considère l'équation différentielle :

$$(H) \quad y'' + py' + qy = 0\,.$$

1) Montrer qu'il existe une et une seule solution y_1 de (H) telle que :

$$y_1(0) = 1\,,\ y_1'(0) = 0 \ \text{ et } \ y_1 \text{ est paire}\,.$$

2) Montrer qu'il existe une et une seule solution y_2 de (H) telle que :

$$y_2(0) = 0\,,\ y_2'(0) = 1 \ \text{ et } \ y_2 \text{ est impaire}\,.$$

3) Montrer que (y_1, y_2) est une base de l'espace vectoriel des solutions de (H).

1) *D'après le théorème de Cauchy-Lipschitz, il existe une et une seule solution de (H) qui vérifie les conditions initiales :*

$$y_1(0) = 1 \quad \text{et} \quad y_1'(0) = 0\,.$$

Montrons que y_1 est paire.
Si on considère l'application $f : x \mapsto y_1(-x)$, f est de classe $\mathscr{C}^2$ et, pour tout $x \in \mathbb{R}$ on a, compte tenu des hypothèses faites sur p et q :

$$\begin{aligned}(f'' + pf' + q)(x) &= y_1''(-x) - p(x)y_1'(-x) + q(x)y_1(-x)\\ &= y_1''(-x) + p(-x)y_1'(-x) + q(-x)y_1(-x)\\ &= (y_1'' + py_1' + q)(-x) = 0\,,\end{aligned}$$

donc f est aussi solution de (H). Puisque $f'(0) = -y_1'(0) = 0$ et $f(0) = y_1(0) = 1$, il résulte de l'unicité dans le théorème de Cauchy-Lipschitz que $f = y_1$, c'est-à-dire y_1 est paire.

2) *Cette question se résout suivant le même principe.*

3) *Supposons qu'il existe des scalaires λ_1 et λ_2 tels que $\lambda_1 y_1 + \lambda_2 y_2 = 0$. En dérivant cette relation puis en prenant la valeur en 0 on aura :*

$$\begin{cases} \lambda_1 y_1(0) + \lambda_2 y_2(0) = 0 \\ \lambda_1 y_1'(0) + \lambda_2 y_2'(0) = 0 \end{cases} \quad \text{soit} \quad \lambda_1 = \lambda_2 = 0\,.$$

La famille (y_1,y_2) est donc libre, et puisque l'on sait d'après le cours que l'espace vectoriel des solutions de (H) est de dimension 2 (c'est encore une conséquence du théorème de Cauchy-Lipschitz !), c'en est une base.

SF17.14 Utiliser le wronskien

L'utilisation du wronskien peut être utile afin de vérifier que n solutions d'un système différentiel d'ordre n sont indépendantes. Néanmoins, dans le cas $n = 2$ (qui est celui que vous rencontrerez pratiquement toujours), il n'a guère d'intérêt pratique, car il est facile de vérifier si deux fonctions données sont ou non proportionnelles !
Son intérêt apparaît donc surtout dans certains exercices théoriques.

Exemple

Soit p continue sur $\mathbb{R}_+$, à valeurs dans $\mathbb{R}$, et intégrable sur $\mathbb{R}_+$.
On considère l'équation différentielle (E) : $y'' + p(x)y = 0$.

1) Montrer que, si y est une solution bornée de (E), alors $\lim\limits_{x\to+\infty} y'(x) = 0$.

2) Montrer que l'ensemble des solutions bornées de (E) est un sous-espace vectoriel de $\mathscr{C}(\mathbb{R}_+,\mathbb{R})$, et préciser sa dimension.

3) Montrer que (E) admet des solutions non bornées.

1) *Soit y une solution bornée de (E). La relation $y'' = -p(x)y$ donne en intégrant :*

$$\forall\, x \in \mathbb{R}_+,\ y'(x) = y'(0) - \int_0^x p(t)y(t)\,\mathrm{d}t \quad (*)$$

Puisque y est bornéee, on a $|p(t)y(t)| \leqslant \|y\|_\infty |p(t)|$ pour tout t donc par comparaison on en déduit que py est intégrable sur $\mathbb{R}_+$.
La relation $(*)$ implique alors que $\lim\limits_{x\to+\infty} y'(x) = \ell$ existe.
Si l'on avait par exemple $\ell > 0$, alors il existerait x_0 tel que pour $x \geqslant x_0$ on ait $y'(x) \geqslant \dfrac{\ell}{2}$ d'où $y(x) - y(x_0) \geqslant \dfrac{\ell}{2}(x - x_0)$ puis $\lim\limits_{x\to+\infty} y(x) = +\infty$: contradiction.
De même, $\ell < 0$ est impossible, et finalement $\ell = 0$.

2) *Soit F l'ensemble des solutions bornées sur $\mathbb{R}$.*
Le fait que F est un sous-espace vectoriel de $\mathscr{C}(\mathbb{R}_+,\mathbb{R})$ est immédiat.
Soient alors y_1, $y_2 \in F$. $y_1'' = -py_1$ et $y_2'' = -py_2$ d'où $y_1''y_2 - y_2''y_1 = 0$ puis $y_1'y_2 - y_2'y_1 = cste$.
Or $\lim\limits_{x\to+\infty} y_1'(x) = \lim\limits_{x\to+\infty} y_2'(x) = 0$ d'après la question précédente et y_1, y_2 sont bornées donc $\lim\limits_{x\to+\infty} (y_1'y_2 - y_2'y_1)(x) = 0$ d'où $cste = 0$.
Ainsi le wronskien de y_1 et y_2 est nul, ce qui implique que la famille (y_1,y_2) est liée. En conclusion : $\dim F \leqslant 1$.

3) *L'ensemble des solutions étant un un espace vectoriel de dimension 2 d'après le théorème de Cauchy-Lipschitz, il existe donc forcément des solutions non bornées.*

■ Comment utiliser les équations différentielles ?

Les équations différentielles interviennent dans de nombreux problèmes. Nous avons déjà eu l'occasion des les utiliser :

- pour chercher le noyau de certaines applications linéaires ;
- pour déterminer les éléments propres de certains endomorphismes ;
- pour trouver la valeur de la somme d'une série entière ;
- pour trouver la valeur d'une intégrale à paramètre.

SF17.15 Dériver une relation

On peut aussi utiliser les équations différentielles pour résoudre certaines équations fonctionnelles, en essayant de dériver la relation proposée.
Nous avons déjà eu l'occasion d'utiliser cette méthode pour résoudre une équation intégrale (voir **[SF10.17]** et exercice **10.7**).
Donnons un autre exemple d'utilisation.

Exemple

Trouver toutes les applications $f : \mathbb{R} \to \mathbb{R}$ dérivables en 0 telles que :

$$\forall\,(x,\,y) \in \mathbb{R}^2,\ f(x+y) = \mathrm{e}^x f(y) + \mathrm{e}^y f(x)\,.$$

- *Soit f une solution au problème, s'il en existe.*
  Pour $x = y = 0$ on obtient $f(0) = 0$.
  De plus, pour tout x réel :

$$\frac{f(x+h)-f(x)}{h} = \frac{\mathrm{e}^x f(h) + \mathrm{e}^h f(x) - f(x)}{h} = \mathrm{e}^x \frac{f(h)-f(0)}{h} + \frac{\mathrm{e}^h - 1}{h} f(x)$$

  donc :

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \mathrm{e}^x f'(0) + f(x)\,.$$

  Il en résulte que f est dérivable sur $\mathbb{R}$ et que :

$$\forall\, x \in \mathbb{R},\ f'(x) = f'(0)\mathrm{e}^x + f(x)\,.$$

  f est donc solution d'une équation différentielle de la forme $y' = y + C\mathrm{e}^x$.
  Il s'agit d'une équation différentielle linéaire du 1er ordre, à coefficients constants, qui se résout aisément. Compte tenu de la condition initiale $f(0) = 0$ on trouve :

$$f(x) = Cx\mathrm{e}^x \quad \text{avec } C \in \mathbb{R}\,.$$

- *On vérifie ensuite facilement que, réciproquement, toutes ces fonctions conviennent.*

SF17.16 Faire apparaître une équation différentielle

Dans de nombreux problèmes où apparaissent des dérivées, il peut être intéressant de considérer les fonctions comme solutions d'une certaine équation différentielle, afin d'utiliser les techniques vues auparavant.
C'est ce que nous avons fait dans le second exemple de **[SF17.12]**. Donnons un autre exemple.

Exemple

Soient α un nombre complexe de partie réelle strictement positive et f une application de classe $\mathscr{C}^1$ de $\mathbb{R}$ dans $\mathbb{C}$ telle que $\lim\limits_{x\to+\infty} f'(x)+\alpha f(x)=0$.

1) Montrer que $\lim\limits_{x\to+\infty} f(x)=0$.

2) En déduire que, si $g\in\mathscr{C}^2(\mathbb{R},\mathbb{R})$ vérifie :

$$\lim_{x\to+\infty} g''(x)+g'(x)+g(x)=0\,,$$

alors $\lim\limits_{x\to+\infty} g(x)=0$.

1) *Posons $g=f'+\alpha f$. La fonction f est ainsi solution de l'équation différentielle : $y'+\alpha y=g$., et d'après le cours, il existe $\lambda\in\mathbb{C}$ tel que :*

$$f(x)=\lambda \mathrm{e}^{-\alpha x}+\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t$$

✓ *Si vous ne vous souvenez plus de ce résultat, il est très facile de le retrouver en utilisant la méthode de variation de la constante, puisqu'une solution de l'équation homogène est $x\mapsto \mathrm{e}^{-\alpha x}$.*

- *Puisque $\alpha=a+\mathrm{i}b$ avec $a>0$ et $b\in\mathbb{R}$, $\mathrm{e}^{-\alpha x}=\mathrm{e}^{-ax}\mathrm{e}^{-\mathrm{i}bx}$ donc $\lim\limits_{x\to+\infty}\mathrm{e}^{-\alpha x}=0$ comme produit d'une fonction qui tend vers 0 par une fonction bornée.*
- *Étudions maintenant la limite du terme $\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t$. Pour démontrer que cette limite est nulle, nous allons employer une méthode qui s'apparente à celle de la démonstration du théorème de Cesàro.*

  Soit $\varepsilon>0$. Puisque $\lim\limits_{t\to+\infty} g(t)=0$ par hypothèse, par définition de la limite :

  $$\exists\, A\geqslant 0 \text{ tel que } \forall\, t\geqslant A,\ |g(t)|\leqslant\varepsilon\,.$$

  On a alors pour tout $x\geqslant A$:

  $$\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t=\mathrm{e}^{-\alpha x}\left(\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t+\int_A^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right),$$

  d'où :

  $$\left|\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right|\leqslant\left|\mathrm{e}^{-\alpha x}\right|\left|\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right|+\left|\mathrm{e}^{-\alpha x}\right|\int_A^x\left|\mathrm{e}^{\alpha t}\right||g(t)|\,\mathrm{d}t\,.$$

  En écrivant toujours $\alpha=a+\mathrm{i}b$ avec $a>0$ et $b\in\mathbb{R}$, puisque $\left|\mathrm{e}^{\mathrm{i}bt}\right|=1$ on obtient :

$$\left|\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| \leqslant \mathrm{e}^{-ax}\left|\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| + \mathrm{e}^{-ax}\int_A^x \mathrm{e}^{at}\,|g(t)|\,\mathrm{d}t$$
$$\leqslant \mathrm{e}^{-ax}\left|\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| + \varepsilon\mathrm{e}^{-ax}\underbrace{\int_A^x \mathrm{e}^{at}\,\mathrm{d}t}_{=\frac{1}{a}(\mathrm{e}^{ax}-\mathrm{e}^{aA})}$$
$$\leqslant \mathrm{e}^{-ax}\left|\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| + \frac{\varepsilon}{a}\cdot$$

Puisque A est fixé et puisque $a>0$, , on a $\lim\limits_{x\to+\infty} \mathrm{e}^{-ax}\left|\int_0^A \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| = 0$ *donc pour x assez grand on aura :*

$$\left|\mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t\right| \leqslant \varepsilon + \frac{\varepsilon}{a}\cdot$$

Ainsi $\lim\limits_{x\to+\infty} \mathrm{e}^{-\alpha x}\int_0^x \mathrm{e}^{\alpha t}g(t)\,\mathrm{d}t = 0$ *puis finalement* $\lim\limits_{x\to+\infty} f(x) = 0$.

2) *En notant comme d'habitude* $\mathrm{j} = \mathrm{e}^{\frac{2\mathrm{i}\pi}{3}}$ *une racine cubique de 1, posons ici* $f = g' - \mathrm{j}^2 g$.
Compte tenu de $1+\mathrm{j}+\mathrm{j}^2 = 0$, *l'hypothèse sur g se traduit par* $\lim\limits_{x\to+\infty} f'(x) - \mathrm{j}f(x) = 0$.
Pour conclure, il suffit d'appliquer deux fois le résultat précédent, ce qui est possible puisque les parties réelles de $-\mathrm{j}$ et de $-\mathrm{j}^2$ sont strictement positives.

À vous de jouer !

Exercice 17.1 **SF13.16** **SF17.1**

Résoudre sur $\mathbb{R}$ l'équation différentielle : $(\mathrm{e}^x - 1)y' + \mathrm{e}^x y = 1$.

Exercice 17.2 **SF17.1** **SF17.11** **SF17.13**

On considère l'équation différentielle :

$$(E) \qquad y' + ay = f(x)$$

où a est un nombre complexe et f une fonction continue de $\mathbb{R}$ dans $\mathbb{C}$, 2π-périodique.

1) Montrer que, si a n'appartient pas à l'ensemble $\mathrm{i}\mathbb{Z}$, il existe une et une seule solution de E qui est 2π-périodique.

2) Lorsque $a \in \mathrm{i}\mathbb{Z}$, montrer que (E) possède une solution 2π-périodique si et seulement si $\int_0^{2\pi} \mathrm{e}^{ax} f(x)\,\mathrm{d}x = 0$.

Exercice 17.3 **SF17.13**

On munit $\mathcal{M}_{n,1}(\mathbb{R})$ de sa structure euclidienne canonique. Soit $A \in \mathcal{M}_n(\mathbb{R})$.
Montrer que les propriétés suivantes sont équivalentes :

(i) A est antisymétrique ;

(ii) chaque solution du système différentiel $X' = AX$ est de norme constante.

Exercice 17.4 **SF17.3** **SF17.6**

On considère l'équation différentielle :

$$(E) \quad 2xy'' + y' - y = 2x + \frac{1}{\operatorname{sh}(\sqrt{2x})} ,$$

d'inconnue $y : \mathbb{R}_+^* \to \mathbb{R}$.
Effectuer un changement de variables $x = \varphi(t)$ dans l'équation différentielle (E) de façon à ce qu'elle devienne une équation à coefficients constants, puis la résoudre.

Exercice 17.5 **SF17.4**

Trouver toutes les solutions sur $\mathbb{R}$ de l'équation différentielle :

$$(x+1)y'' - y' - xy = \mathrm{e}^{-x} \quad (E)$$

(on pourra remarquer que $x \mapsto \mathrm{e}^x$ est solution de l'équation homogène).

Exercice 17.6 **SF17.5**

Résoudre sur $]0\,;1[$ l'équation :

$$x(1-x)y'' + (1-3x)y' - y = 0$$

en commençant par rechercher une solution développable en série entière.

Exercice 17.7 **SF17.7**

On pose (E) : $\quad x^2y'' + 4xy' + (2-x^2)y = 1$.
Résoudre cette équation différentielle en posant $z = x^2y$.
Déterminer, s'il en existe, les solutions définies sur $\mathbb{R}$.

Exercice 17.8 **SF17.10**

Soit n un entier $\geqslant 2$. Soient U, $V \in \mathcal{M}_{n,1}(\mathbb{R})$, non nuls, et $B = U{}^{\mathrm{t}}V \in \mathcal{M}_n(\mathbb{R})$.
Étant donné $\lambda \in \mathbb{R}$, on considère la matrice :

$$A = \lambda \mathrm{I}_n + B \ \in \mathcal{M}_n(\mathbb{R}) ,$$

et le système différentiel $X' = AX$.

1) Exprimer la solution $t \mapsto X(t)$ telle que $X(0) = X_0$, où $X_0 \in \mathcal{M}_{n,1}(\mathbb{R})$ est donné.

2) À quelle condition toute solution X a-t-elle une limite quand $t \to +\infty$?

Exercice 17.9 **SF17.15**

1) Déterminer l'ensemble des applications de classe $\mathscr{C}^2$ sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ telles que :

$$\forall\,(x,y)\in\mathbb{R}^2,\ f(x+y)+f(x-y)=2f(x)f(y)\,.$$

2) Même question en supposant seulement f continue.

Exercice 17.10 **SF17.7** **SF17.13**

1) Soit I un intervalle de $\mathbb{R}$, $a\colon I\to\mathbb{R}$ de classe $\mathscr{C}^1$ et $b\colon I\to\mathbb{R}$ continue.
Trouver une condition nécessaire et suffisante pour que l'équation différentielle :

$$y''+a(x)y'+b(x)y=0$$

admette deux solutions y_1 et y_2 non nulles telles que $y_2=xy_1$.

2) *Application :*
Résoudre l'équation différentielle : $y''+2xy'+(1+x^2)y=0$.

Exercice 17.11 **SF17.13**

Soit $p:\mathbb{R}\to\mathbb{R}^+$ une fonction continue non nulle.

1) Montrer que toute solution sur $\mathbb{R}$ de l'équation différentielle $y''+p(x)y=0$ s'annule en au moins un point.

2) Soit y une solution sur $\mathbb{R}$ de l'équation $y''-p(x)y=0$, autre que la solution nulle.
Montrer que y et y' s'annulent au plus une fois.

Exercice 17.12 **SF17.14**

Soient p et q deux fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$.
On considère l'équation différentielle : $(E)\ y''+py'+qy=0$.

1) Soit y une solution non nulle de (E). Montrer que les zéros de y sont isolés.

2) Soient y_1 et y_2 deux solutions linéairement indépendantes de (E).

a) Montrer que y_1 et y_2 ne peuvent pas avoir de zéro commun.

b) Montrer que si a et b sont deux zéros consécutifs de y_1 $(a<b)$, il existe un et un seul $c\in\,]a\,;b[$ tel que $y_2(c)=0$.

■ Comment étudier la limite d'une fonction de plusieurs variables ?

SF18.1 Revenir à la définition

Pour montrer que la limite d'une fonction f quand x tend vers a, c'est-à-dire quand $\|x-a\|$ tend vers 0, est égale à ℓ, on peut former l'expression $\|f(x)-\ell\|$ et essayer de la majorer.
Puisque toutes les normes sur un espace de dimension finie sont équivalentes, il faut en général utiliser une norme adaptée.

Exemple

Soit $f : \begin{cases} \mathbb{R}^2 \setminus \{(0,0)\} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \dfrac{\sin(xy)}{|x|+|y|} \end{cases}$. Montrer que $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$.

Il est commode ici d'utiliser la norme : $\|(x,y)\|_\infty = \max\{|x|, |y|\}$.
En effet, pour tout $(x,y) \in \mathbb{R}^2$ *on a :*

$$|\sin(xy)| \leqslant |xy| \leqslant \|(x,y)\|_\infty^2 \quad \text{et} \quad |x|+|y| \geqslant \|(x,y)\|_\infty ,$$

d'où :

$$\forall\, (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\},\ |f(x,y)| \leqslant \|(x,y)\|_\infty$$

puis la limite demandée.

SF18.2 Utiliser la limite d'une fonction composée

De façon simplifiée : si $\lim\limits_{x\to a} g(x) = b$ et $\lim\limits_{y\to b} f(y) = \ell$, alors $\lim\limits_{x\to a} f\circ g(x) = \ell$.
Par conséquent, si l'on peut trouver des fonctions g_1 et g_2 telles que

$$\lim_{x\to a} g_1(x) = \lim_{x\to a} g_2(x) = b \quad \text{et} \quad \lim_{x\to a} f\circ g_1(x) \neq \lim_{x\to a} f\circ g_2(x),$$

f ne peut admettre de limite en b.

Par exemple, dans $\mathbb{R}^2$, on pourra prendre pour g des fonctions de la forme $x \mapsto (x,\lambda x)$ (limite le long d'une droite), ou de la forme $x \mapsto (x,\lambda x^2)$ (limite le long d'une parabole), ou de la forme $x \mapsto (x^\alpha, x^\beta)$.

Exemple

Montrer que $f : \begin{cases} \mathbb{R}^2 \setminus \{(0,0)\} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \dfrac{|x|\sqrt{|y|}}{x^2+y^2} \end{cases}$ n'a pas de limite en $(0,0)$.

Pour tout $x \neq 0$, $f(x,0) = 0$ *donc* $\lim\limits_{x\to 0} f(x,0) = 0$.

Mais $\lim\limits_{\substack{x\to 0 \\ x\neq 0}} f(x,x^2) = \lim\limits_{x\to 0} \dfrac{1}{1+x^2} = 1 \neq 0$, *d'où le résultat demandé.*

Exemple

Soit $U = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2-2z^2 \neq 0\}$ et

$$f : \begin{cases} U & \longrightarrow & \mathbb{R} \\ (x,y,z) & \longmapsto & \dfrac{x^3+y^3+z^3}{x^2+y^2-2z^2}\,. \end{cases}$$

f a-t-elle une limite en $(0,0,0)$?

- *On remarque déjà que U est bien un ouvert de $\mathbb{R}^3$, car c'est l'image réciproque de l'ouvert $\mathbb{R}^*$ de $\mathbb{R}$ par l'application continue $(x,y,z) \mapsto x^2+y^2-2z^2$.*
- *On a : $\lim\limits_{x\to 0} f(x,0,0) = \lim\limits_{x\to 0} x = 0$. Donc la limite si elle existe ne peut être que 0.*

  Mais pour $x \neq 0$, $f(x,x,x+x^2) = \dfrac{2x^3+x^3(1+x)^3}{2x^2-2x^2(1+x)^2} = \dfrac{3x^3+o(x^3)}{-4x^3+o(x^3)}$,

  donc $\lim\limits_{x\to 0} f(x,x,x+x^2) = -\dfrac{3}{4}$, et f ne peut admettre de limite en $(0,0,0)$.

■ Comment étudier la continuité d'une fonction de plusieurs variables ?

SF18.3 Utiliser les théorèmes généraux

On sait que les applications coordonnées $\varphi_i : (x_1,\dots,x_n) \longmapsto x_i$ sont continues de $\mathbb{R}^n$ dans $\mathbb{R}$ (car linéaires sur un espace de dimension finie).

Il en résulte que toute fonction polynomiale en les coordonnées est continue. À partir de là, en utilisant les théorèmes généraux sur la continuité (produit, composée, quotient...), on obtient la continuité de la plupart des fonctions que vous aurez à considérer là où elles sont définies, sauf peut-être en quelques points litigieux.

SF18.4 Revenir à la définition

Pour étudier la continuité en des points litigieux, on revient à la définition : $\lim\limits_{x\to a} f(x) = f(a)$. On est donc ramené à un calcul de limite.

Exemple

On considère $f : \mathbb{R}^2 \to \mathbb{R}$ définie par : $f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0\,. \end{cases}$

Montrer que f est continue sur $\mathbb{R}^2$.

- *Les applications $(x,y) \mapsto x^2y$ et $(x,y) \mapsto x^2+y^2$ sont continues d'après les théorèmes rappelés ci-dessus, donc f est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$ comme quotient de fonctions continues dont le dénominateur ne s'annule pas.*
- *On a $x^2 \leqslant x^2+y^2$ et $|y| \leqslant \sqrt{x^2+y^2}$ donc $|f(x,y)| \leqslant \|(x,y)\|_2$. Cela prouve que $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$, donc f est aussi continue en $(0,0)$.*

SF18.5 Utiliser les théorèmes généraux

Comme pour les applications de $\mathbb{R}$ dans $\mathbb{R}$, pour montrer qu'une application de $\mathbb{R}^p$ dans $\mathbb{R}^n$ est différentiable, on peut utiliser les théorèmes rappelés en **[S18.5]** et **[S18.8]** : combinaisons linéaires, composée, produit et quotient pour les fonctions à valeurs dans $\mathbb{R}$, etc...

Exemple

Soit E un espace vectoriel euclidien, et $\varphi : \begin{cases} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \|x\|^2 \end{cases}$.

Montrer que φ est différentiable sur E et trouver sa différentielle.

On peut écrire $\varphi = B(f,g)$, où $f = g = \mathrm{Id}_E$ et où B est le produit scalaire sur E.

f et g sont différentiables sur E et pour tout $a \in E$, $\mathrm{d}f_a = f$ et $\mathrm{d}g_a = g$.

✓ *Il s'agit là d'un cas particulier d'un résultat à connaître : si f est une application linéaire d'un espace vectoriel E vers un espace vectoriel F, elle est différentiable sur E et pour tout $a \in E$, $\mathrm{d}f_a = f$.*

B est une application bilinéaire, donc les théorèmes du cours permettent d'affirmer que φ est différentiable en tout point et l'on a :

$$\begin{aligned} \forall\, a \in E,\ \forall\, h \in E,\ \mathrm{d}\varphi_a(h) &= B\big(\mathrm{d}f_a(h), g(a)\big) + B\big(f(a), \mathrm{d}g_a(h)\big) \\ &= \langle h \,|\, a\rangle + \langle a \,|\, h\rangle = 2\,\langle a \,|\, h\rangle \ . \end{aligned}$$

SF18.6 Utiliser les dérivées partielles

Si les dérivées partielles existent et sont continues, la fonction est de classe $\mathscr{C}^1$ et a fortiori différentiable. De plus, on peut alors exprimer la différentielle à l'aide des dérivées partielles.

Exemple

Montrer que l'application $\det : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ M & \longmapsto & \det M \end{cases}$ est de classe $\mathscr{C}^1$ et préciser sa différentielle.

- *Il s'agit ici de calculer les dérivées partielles :*

$$\frac{\partial \det}{\partial E_{ij}}(M) = \lim_{h \to 0} \frac{\det(M + hE_{ij}) - \det M}{h} \ ,$$

où $\big(E_{ij}\big)_{1 \leqslant i,j \leqslant n}$ est la base canonique de $\mathcal{M}_n(\mathbb{R})$. Or en développant le déterminant selon la j^e colonne, on a :

$$\det M = \sum_{k=1}^{n} m_{kj} M_{kj} \ ,$$

où M_{kj} est le cofacteur d'indice (k,j) dans M.

Lorsqu'on ajoute à M la matrice hE_{ij}, les cofacteurs qui interviennent dans le développement du déterminant selon la j^e colonne sont inchangés, et les coefficients m_{kj} sont inchangés sauf m_{ij}, de sorte que :

$$\det(M+hE_{ij}) - \det M = \big((m_{ij}+h) - m_{ij}\big)M_{ij} = hM_{ij}\,,$$

donc la dérivée partielle $\dfrac{\partial \det}{\partial E_{ij}}(M)$ *existe et est égale à* M_{ij}.

De plus, le cofacteur M_{ij} est une fonction polynomiale des coefficients de M, c'est donc une fonction continue de M ; cela prouve que les dérivées partielles sont continues, donc l'application det *est de classe* $\mathscr{C}^1$.

- *De plus, sa différentielle en tout point $M \in \mathcal{M}_n(\mathbb{R})$ est donnée par la formule :*

$$\forall\, H \in \mathcal{M}_n(\mathbb{R}),\ \mathrm{d}(\det)_M(H) = \sum_{1\leqslant i,j\leqslant n} h_{ij}\frac{\partial \det}{\partial E_{ij}}(M) = \sum_{1\leqslant i,j\leqslant n} M_{ij}h_{ij}\,.$$

On reconnaît là le produit scalaire de la matrice H avec la matrice $(M_{ij})_{1\leqslant i,j\leqslant n}$, qui est la comatrice de M (il s'agit du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$).
Ainsi, la différentielle de det *en M est l'application :*

$$H \mapsto \langle \mathrm{com}(M)\,|\,H\rangle = \mathrm{tr}(\widetilde{M}H)\,,$$

où $\widetilde{M} = {}^{\mathrm{t}}\mathrm{com}(M)$ désigne la matrice complémentaire de M.

SF18.7 **Revenir à la définition**

Le calcul de la différentielle de f à partir de la définition consiste à former la différence $f(a+h)-f(a)$ et à y faire apparaître un terme linéaire en h (ce sera $\mathrm{d}f_a(h)$) plus un terme en $o(\|h\|)$.

Exemple

Montrer que l'application $f\colon M \mapsto M^{-1}$ est différentiable sur $\mathrm{GL}_n(\mathbb{R})$, et trouver sa différentielle.

On supposera choisie une norme d'algèbre sur $\mathcal{M}_n(\mathbb{R})$. On rappelle que $\mathrm{GL}_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$ (image réciproque de $\mathbb{R}^*$ par l'application continue det*).*

- *On sait que pour tout $M \in \mathrm{GL}_n(\mathbb{R})$:*

$$M^{-1} = \frac{1}{\det M}\,{}^{\mathrm{t}}\mathrm{com}(M)\,.$$

Or les applications $M \mapsto \det M$ et $M \mapsto {}^{\mathrm{t}}\mathrm{com}(M)$ sont continues car fonctions polynomiales des coefficients de M, donc l'application $M \mapsto M^{-1}$ est continue sur $\mathrm{GL}_n(\mathbb{R})$ (le même argument montre d'ailleurs qu'elle est de classe $\mathscr{C}^1$).

- *Soit M une matrice inversible, et H une matrice de norme assez petite pour que $M+H$ soit encore inversible (c'est possible puisque $\mathrm{GL}_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$).*
On a alors :

$$(M+H)^{-1} - M^{-1} = (M+H)^{-1}\big(\mathrm{I}_n - (M+H)M^{-1}\big) = -(M+H)^{-1}HM^{-1}$$

donc : $(M+H)^{-1} - M^{-1} + M^{-1}HM^{-1} = \big(M^{-1} - (M+H)^{-1}\big)HM^{-1}$.

✓ *L'idée ici est de s'inspirer du calcul suivant dans* $\mathbb{R}$ *:*

$$\frac{1}{m+h} - \frac{1}{m} = -\frac{h}{m(m+h)} \approx -\frac{h}{m^2}\cdot$$

Or :

$$\big\|\big(M^{-1} - (M+H)^{-1}\big)HM^{-1}\big\| \leqslant \big\|M^{-1} - (M+H)^{-1}\big\|\,\|H\|\,\big\|M^{-1}\big\|\ ,$$

et puisque $\lim\limits_{H\to 0_n}\big(M^{-1} - (M+H)^{-1}\big) = 0$ *par continuité de l'application* $M \mapsto M^{-1}$*, on en déduit que :*

$$\big(M^{-1} - (M+H)^{-1}\big)HM^{-1} \underset{H\to 0_n}{=} o(\|H\|)\,.$$

Finalement :

$$(M+H)^{-1} = M^{-1} - M^{-1}HM^{-1} + o(\|H\|)\ ,$$

et puisque l'application $H \mapsto -M^{-1}HM^{-1}$ *est linéaire, cela prouve que l'application* $f\colon M \mapsto M^{-1}$ *est différentiable en tout point* $M \in \mathrm{GL}_n(\mathbb{R})$ *et que :*

$$\forall\, H \in \mathcal{M}_n(\mathbb{R}),\ \mathrm{d}f_M(H) = -M^{-1}HM^{-1}.$$

■ Comment étudier les dérivées partielles en un point ?

SF18.8 Utiliser les théorèmes généraux et la définition

Pour étudier si une fonction de plusieurs variables possède des dérivées partielles, on commence par essayer d'appliquer les théorèmes généraux sur les fonctions de classe $\mathscr{C}^1$.

Aux points qui posent problème, on peut calculer les dérivées partielles en revenant à la définition (limite du taux d'accroissement des application partielles). Pour vérifier ensuite que f est (ou non) de classe $\mathscr{C}^1$, il faut examiner la continuité des dérivées partielles ; on est donc encore ramené à un calcul de limite.

Le même principe s'applique pour étudier les dérivées partielles d'ordre 2 ; mais les calculs sont un peu plus longs…

Exemple

On considère $f\colon \mathbb{R}^2 \to \mathbb{R}$ définie par : $f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0\,. \end{cases}$

f est-elle de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$?

- *Nous avons déjà vu dans un exemple page 638 que* f *est continue sur* $\mathbb{R}^2$*. D'après les théorèmes généraux sur les dérivées d'une fonction d'une variable réelle, les dérivées partielles de* f *existent sur* $\mathbb{R}^2 \setminus \{(0,0)\}$ *et,*

pour tout $(x,y) \neq (0,0)$ *:*

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = \dfrac{2xy(x^2+y^2) - x^2y(2x)}{(x^2+y^2)^2} = \dfrac{2xy^3}{(x^2+y^2)^2}\,, \\ \dfrac{\partial f}{\partial y}(x,y) = \dfrac{x^2(x^2+y^2) - x^2y(2y)}{(x^2+y^2)^2} = \dfrac{x^4-x^2y^2}{(x^2+y^2)^2}\,. \end{cases}$$

Les deux applications $\dfrac{\partial f}{\partial x}$ *et* $\dfrac{\partial f}{\partial y}$ *sont continues sur* $\mathbb{R}^2 \setminus \{(0,0)\}$ *d'après les théorèmes habituels, donc* f *est de classe* $\mathscr{C}^1$ *sur* $\mathbb{R}^2 \setminus \{(0,0)\}$.

- *On calcule les dérivées partielles en* $(0,0)$ *en revenant à la définition :*

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} 0 = 0\,,$$

donc $\dfrac{\partial f}{\partial x}(0,0)$ *existe et est égale à* 0.

On trouve de la même façon que $\dfrac{\partial f}{\partial y}(0,0)$ *existe et vaut* 0.

Cependant, $\dfrac{\partial f}{\partial x}$ *n'est pas continue en* $(0,0)$ *puisque, par exemple :*

$$\lim_{x\to 0} \frac{\partial f}{\partial x}(x,0) = 0 \quad \text{et} \quad \lim_{x\to 0} \frac{\partial f}{\partial x}(x,x) = \lim_{x\to 0} \frac{2x^4}{4x^4} = \frac{1}{2}\,.$$

Ainsi, f *n'est pas de classe* $\mathscr{C}^1$ *sur* $\mathbb{R}^2$.

Exemple

Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ définie par : $f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0\,. \end{cases}$

f est-elle de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$?

- f *est de classe* $\mathscr{C}^2$ *sur* $\mathbb{R}^2 \setminus \{(0,0)\}$, *comme quotient de fonctions polynomiales en les variables* x *et* y, *dont le dénominateur ne s'annule pas.*
- *Pour* $(x,y) \neq (0,0)$ *on calcule :*

$$\frac{\partial f}{\partial x}(x,y) = \frac{(3x^2y-y^3)(x^2+y^2) - 2x^2y(x^2-y^2)}{(x^2+y^2)^2} = \frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2}\,.$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2}\,.$$

Puisque $\dfrac{f(x,0)-f(0,0)}{x} = 0$ *et* $\dfrac{f(0,y)-f(0,0)}{y} = 0$, *les dérivées partielles de* f *en* $(0,0)$ *existent et sont égales à* 0.

En posant $x = r\cos\theta$ *et* $y = r\sin\theta$, *on a*

$$\left|\frac{\partial f}{\partial x}(x,y)\right| \leqslant 6r \quad \text{et} \quad \left|\frac{\partial f}{\partial y}(x,y)\right| \leqslant 6r\,,$$

ce qui montre que $\lim\limits_{(x,y)\to(0,0)} \dfrac{\partial f}{\partial x}(x,y) = 0 = \dfrac{\partial f}{\partial x}(0,0)$, *donc* $\dfrac{\partial f}{\partial x}$ *est continue sur* $\mathbb{R}^2$, *et il en est de même de* $\dfrac{\partial f}{\partial y}$. *Donc* f *est de classe* $\mathscr{C}^1$ *sur* $\mathbb{R}^2$.

- *On calcule ensuite :*
$$\lim_{x\to 0}\frac{\dfrac{\partial f}{\partial y}(x,0)-\dfrac{\partial f}{\partial y}(0,0)}{x}=\lim_{x\to 0}\frac{x^5}{x^5}=1$$
donc $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)=\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right)(0,0)$ *existe et est égale à* 1. *De même :*
$$\lim_{y\to 0}\frac{\dfrac{\partial f}{\partial x}(0,y)-\dfrac{\partial f}{\partial x}(0,0)}{x}=\lim_{y\to 0}\frac{-y^5}{y^5}=-1$$
donc $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)=\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right)(0,0)$ *existe et est égale à* -1.
Puisque $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)\neq\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$, *il résulte du théorème de Schwarz, par contraposition, que* f *n'est pas de classe* $\mathscr{C}^2$ *sur* $\mathbb{R}^2$.

SF18.9 Utiliser la différentielle

Si f est de classe $\mathscr{C}^1$ sur $U\subset\mathbb{R}^p$, on sait que f admet en tout point a de U une différentielle $\mathrm{d}f_a$ qui est donnée par :
$$\forall\,(h_1,\dots,h_p)\in\mathbb{R}^p,\ \mathrm{d}f_a(h)=\sum_{j=1}^{p}h_j\frac{\partial f}{\partial x_j}(a)\,.$$
On sait de plus que, dans ce cas, pour tout vecteur $\overrightarrow{v}$ de E, f admet une dérivée en a selon $\overrightarrow{v}$ donnée par : $D_{\overrightarrow{v}}f(a)=\mathrm{d}f_a(\overrightarrow{v})$.
Ces propriétés peuvent être utilisées pour montrer, par l'absurde, que f n'est pas de classe $\mathscr{C}^1$ en un certain point a.
Mais cela peut être aussi utilisé pour montrer que f est de classe $\mathscr{C}^1$, car si $\mathrm{d}f_a$ existe en tout point et si l'application $a\mapsto\mathrm{d}f_a$ est continue, les dérivées partielles de f existent et sont alors continues.

Exemple

Montrer que l'application $f\colon\mathbb{R}^2\to\mathbb{R}$ définie par :
$$f(x,y)=\begin{cases}\dfrac{x^3}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } x=y=0\end{cases}$$
est continue sur $\mathbb{R}^2$, admet en $(0,0)$ des dérivées selon tout vecteur, mais n'est pas de classe $\mathscr{C}^1$.

- f *est continue sur* $\mathbb{R}^2\setminus\{(0,0)\}$ *par les théorèmes usuels. Puisque* $|f(x,y)|\leqslant|x|$ *pour tout* $(x,y)\in\mathbb{R}^2$, *on a* $\lim\limits_{(x,y)\to(0,0)}f(x,y)=0=f(0,0)$, *donc* f *est aussi continue en* $(0,0)$.
- *Pour tout vecteur* $\overrightarrow{v}=(v_1,v_2)$ *non nul on a*
$$\lim_{t\to 0}\frac{f((0,0)+t\overrightarrow{v})-f(0,0)}{t}=\lim_{t\to 0}\frac{t^3v_1^3}{t^3(v_1^2+v_2^2)}=\frac{v_1^3}{v_1^2+v_2^2}\,,$$
donc f *admet bien en* $(0,0)$ *une dérivée selon le vecteur* $\overrightarrow{v}$, *donnée par*
$$\frac{\partial f}{\partial\overrightarrow{v}}(0.0)=\frac{v_1^3}{v_1^2+v_2^2}\,.$$

- *En particulier, pour* $\overrightarrow{v} = (1,0)$ *puis* $\overrightarrow{v} = (0,1)$, *on trouve qu'elle admet en* $(0,0)$ *des dérivées partielles :*

$$\frac{\partial f}{\partial x}(0,0) = 1 \quad \text{et} \quad \frac{\partial f}{\partial y}(0,0) = 0\,.$$

Si f *était de classe* $\mathscr{C}^1$, *elle admettrait en* $(0,0)$ *une différentielle donnée par :*

$$\mathrm{d}f_{(0,0)} : (h_1,h_2) \mapsto h_1\frac{\partial f}{\partial x}(0,0) + h_2\frac{\partial f}{\partial y}(0,0) = h_1\,,$$

et on devrait avoir :

$$\frac{\partial f}{\partial \overrightarrow{v}}(0,0) = \mathrm{d}f_{(0,0)}(\overrightarrow{v}) = v_1\,,$$

ce qui est en contradiction avec le résultat obtenu précédemment.

Exemple

Montrer que l'application $f : \left\{\begin{array}{ccc} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ M & \longmapsto & {}^{\mathrm{t}}MM \end{array}\right.$ est de classe $\mathscr{C}^1$.

Pour toutes matrices $M,\ H \in \mathcal{M}_n(\mathbb{R})$ *on a :*

$$f(M+H) - f(M) = {}^{\mathrm{t}}HM + {}^{\mathrm{t}}MH + {}^{\mathrm{t}}HH\,.$$

Munissons $\mathcal{M}_n(\mathbb{R})$ *de la norme* $\|\ \|_\infty$ *définie par :*

$$\forall\, A = (a_{ij})_{1\leqslant i,j\leqslant n},\ \|A\|_\infty = \max\left\{|a_{ij}|\,,\ (i,j) \in [\![1\,;n]\!]^2\right\}\,.$$

Pour tout $(i,j) \in [\![1\,;n]\!]^2$ *on a, si* $H = (h_{ij})_{1\leqslant i,j\leqslant n}$ *:*

$$({}^{\mathrm{t}}HH)_{ij} = \sum_{k=1}^{n} h_{ki}h_{kj}$$

donc

$$\left|({}^{\mathrm{t}}HH)_{ij}\right| \leqslant \sum_{k=1}^{n} |h_{ki}h_{kj}| \leqslant n\,\|H\|_\infty^2$$

d'où : $\|{}^{\mathrm{t}}HH\|_\infty \leqslant n\,\|H\|_\infty^2$.
On peut donc écrire :

$$f(M+H) - f(M) = {}^{\mathrm{t}}HM + {}^{\mathrm{t}}MH + o\Big(\|H\|_\infty\Big)$$

et, puisque l'application $H \mapsto {}^{\mathrm{t}}HM + {}^{\mathrm{t}}MH$ *est linéaire, on reconnaît là un développement limité d'ordre* 1.
Ainsi, f *est différentiable en tout point* $M \in \mathcal{M}_n(\mathbb{R})$ *et*

$$\forall\, H \in \mathcal{M}_n(\mathbb{R}),\ \mathrm{d}f_M(H) = {}^{\mathrm{t}}HM + {}^{\mathrm{t}}MH\,.$$

L'application $M \mapsto \mathrm{d}f_M$ *étant linéaire sur un espace vectoriel de dimension finie, elle est continue, ce qui prouve que* f *est de classe* $\mathscr{C}^1$.

SF18.10 Se ramener à une équation plus simple par un changement de variables

Nous ne considérons ici que des fonctions de deux variables, mais les méthodes décrites se généralisent sans problème aux fonctions de n variables ; seules les écritures sont plus lourdes.

- On sait résoudre certaines équation aux dérivées partielles :
  - $\dfrac{\partial f}{\partial x}(x,y) = 0$: on utilise alors le résultat suivant.

    Soit f une application de classe $\mathscr{C}^1$ définie sur un *ouvert convexe* U de $\mathbb{R}^2$ et à valeurs dans $\mathbb{R}$.

    Si $\dfrac{\partial f}{\partial x} = 0$, alors il existe un intervalle ouvert I de $\mathbb{R}$ et une fonction g de classe $\mathscr{C}^1$ sur I tels que :

    $$\forall\,(x,y) \in U\ ,\ f(x,y) = g(y)\,.$$

  - $\dfrac{\partial f}{\partial x}(x,y) = h(x,y)\,.$

    On se ramène au cas précédent en considérant une primitive de la fonction $x \mapsto h(x,y)$; on obtient que f est de la forme :

    $$f(x) = g(y) + \int_a^x h(t,y)\,\mathrm{d}t\,.$$

  - $\dfrac{\partial f}{\partial x}(x,y) + a(x,y)f(x,y) = h(x,y)\,.$

    On considère alors cette équation comme une équation différentielle linéaire du premier ordre en la 1re application partielle. Dans la résolution de cette équation différentielle, les « constantes » seront en fait des constantes « par rapport à x », c'est-à-dire des fonctions de y.

  Ces résultats concernent des équations de la forme $\dfrac{\partial f}{\partial x} = \cdots$, mais nous avons bien sûr des résultats analogues pour les équations de la forme $\dfrac{\partial f}{\partial y} = \cdots$.

- Dans le cas d'une équation aux dérivées partielles du premier ordre plus générale, on essaie de se ramener aux cas précédents à l'aide d'un changement de variables, généralement fourni par l'énoncé. Ce changement de variables doit être une bijection de classe $\mathscr{C}^1$ entre deux ouverts.

  Si ce changement de variables permet d'exprimer x et y en fonction de u et v (et réciproquement), on posera $F(u,v) = f(x,y)$ comme nouvelle fonction inconnue. On peut alors exprimer les dérivées partielles de f en fonction de celles de F (ou le contraire) à l'aide des formules suivantes (où les notations ont été abrégées : on a omis les (x,y) et (u,v)) :

$$\begin{cases} \dfrac{\partial F}{\partial u} &= \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial u} \\ \dfrac{\partial F}{\partial v} &= \dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial v} \end{cases} \quad \text{ou} \quad \begin{cases} \dfrac{\partial f}{\partial x} &= \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial x} \\ \dfrac{\partial f}{\partial y} &= \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial y} \end{cases}.$$

Exemple

Soit $U = \{(x,y) \in \mathbb{R}^2 \mid x+y>0\}$.
En utilisant les coordonnées polaires, déterminer les $f \in \mathscr{C}^1(U,\mathbb{R})$ telles que :

$$\forall\,(x,y) \in U,\ x\frac{\partial f}{\partial x}(x,y) + y\frac{\partial f}{\partial y}(x,y) = \sqrt{x+y}\,.$$

- *Déjà, U est bien un ouvert de $\mathbb{R}^2$, puisque c'est l'image réciproque de $]0\,;+\infty[$ par l'application continue $(x,y) \mapsto x+y$.*
- *Si l'on pose $D = \mathbb{R}_+^* \times \left]-\frac{\pi}{4}\,;\frac{3\pi}{4}\right[$, l'application*

$$\varphi\colon (r,\theta) \mapsto (x,y) = (r\cos\theta, r\sin\theta)$$

est bijective de D sur U et de classe $\mathscr{C}^1$.
Pour préciser φ^{-1}, on remarque que si $\theta \in \left]-\frac{\pi}{4}\,;\frac{3\pi}{4}\right[$ alors $\theta - \frac{\pi}{4} \in \left]-\frac{\pi}{2}\,;\frac{\pi}{2}\right[$ donc :

$$\theta = \frac{\pi}{4} + \operatorname{Arctan}\left(\tan\left(\theta - \frac{\pi}{4}\right)\right) = \frac{\pi}{4} + \operatorname{Arctan}\left(\frac{\tan\theta - 1}{\tan\theta + 1}\right)$$
$$= \frac{\pi}{4} + \operatorname{Arctan}\left(\frac{y-x}{y+x}\right).$$

Donc φ^{-1} est l'application de U dans D définie par :

$$\varphi^{-1}\colon (x,y) \mapsto \left(\sqrt{x^2+y^2}\,,\frac{\pi}{4} + \operatorname{Arctan}\left(\frac{y-x}{y+x}\right)\right).$$

Elle est bien de classe $\mathscr{C}^1$ d'après les théorèmes usuels.

- *On pose alors $F = f \circ \varphi$, de sorte que F est de classe $\mathscr{C}^1$ sur D et $\forall\,(r,\theta) \in D,\ F(r,\theta) = f(r\cos\theta, r\sin\theta)$.*
On a alors :

$$\begin{cases} \dfrac{\partial F}{\partial r} = \cos\theta\dfrac{\partial f}{\partial x} + \sin\theta\dfrac{\partial f}{\partial y} \\ \dfrac{\partial F}{\partial \theta} = -r\sin\theta\dfrac{\partial f}{\partial x} + r\cos\theta\dfrac{\partial f}{\partial y} \end{cases}$$

donc $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = r\dfrac{\partial F}{\partial r}$.
L'équation proposée est donc équivalente à $r\dfrac{\partial F}{\partial r} = \sqrt{r\cos\theta + r\sin\theta}$ soit $\dfrac{\partial F}{\partial r} = \dfrac{1}{\sqrt{r}}\sqrt{\cos\theta + \sin\theta}$.
- *Ses solutions sont données par :*

$$F(r,\theta) = 2\sqrt{r}\sqrt{\cos\theta + \sin\theta} + g(\theta),$$

où g est une application quelconque de classe $\mathscr{C}^1$ sur $\left]-\frac{\pi}{4}\,;\frac{3\pi}{4}\right[$.
En utilisant φ^{-1}, on obtient que les solutions de l'équation initiale sont les applications de la forme :

$$f(x,y) = 2\sqrt{x+y} + h\left(\frac{y-x}{y+x}\right),$$

où h est une application quelconque de classe $\mathscr{C}^1$ sur $\mathbb{R}$.

■ Comment étudier les extrema d'une fonction de plusieurs variables ?

La méthode utilisée diffère sensiblement suivant que l'on recherche les extrema sur une partie ouverte ou non.

SF18.11 Utiliser le théorème du cours

Lorsque f est une application de classe $\mathscr{C}^1$ définie sur un *ouvert* U de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^n$, les extrema de f sont à chercher parmi les points critiques, c'est-à-dire les points de U où le gradient de f s'annule.
Pour chaque point critique a, il faut ensuite étudier le signe de $f(x) - f(a)$ au voisinage de a, pour vérifier s'il s'agit bien d'un extremum local.

Exemple

1) Déterminer les extrema locaux de la fonction f définie sur $\mathbb{R}^2$ par :

$$\forall\,(x,y) \in \mathbb{R}^2,\ f(x,y) = x^4 + y^4 - 4(x-y)^2\,.$$

2) Montrer que f possède un minimum global sur $\mathbb{R}^2$. En quels points ?

1) *f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ car c'est une fonction polynomiale en x et y. De plus, $\mathbb{R}^2$ est un ouvert, donc les extrema de f sont à chercher parmi ses points critiques. Pour les trouver, on résout :*

$$\nabla f(x,y) = 0 \iff \begin{cases} 4x^3 - 8(x-y) = 0 \\ 4y^3 + 8(x-y) = 0 \end{cases}$$
$$\iff (x,y) \in \{(0,0),(-2,2),(2,-2)\}\,.$$

- Étude en $(0,0)$

  On remarque que $f(x,x) = 2x^4 \geqslant 0 = f(0,0)$ pour tout x, mais $f(x,-x) = 2x^4 - 16x^2 = 2x^2(x^2-8)$ est strictement négatif pour $0 < x^2 < 8$. Donc f ne garde pas un signe constant au voisinage de $(0,0)$, c'est-à-dire prend localement à la fois des valeurs supérieures et inférieures à $f(0,0)$: ce n'est pas un extremum local.

- Étude en $(2,-2)$

  Pour étudier le comportement de f au voisinage de $(2,-2)$, on étudie le signe de $f(2+h,-2+k) - f(2,-2)$ pour h, k « petits » :

$$\begin{aligned} f(2+h,-2+k) - f(2,-2) &= (2+h)^4 + (-2+k)^4 - 4(4+h-k)^2 + 32 \\ &= 20h^2 + 8hk + 20k^2 + 8h^3 - 8k^3 + h^4 + k^4\,. \end{aligned}$$

En écrivant $h = r\cos\theta$ et $k = r\sin\theta$ on obtient, pour $r \neq 0$:

$$\frac{f(2+h,-2+k)-f(2,-2)}{r^2} \underset{r\to 0}{=} 20 + 8\cos\theta\sin\theta + o(1)$$

ce qui montre que, localement, la différence est du signe de $20 + 8\cos\theta\sin\theta$, donc est positive. Il s'agit donc d'un minimum local.

- Étude en $(-2,2)$

  Puisque $f(y,x) = f(x,y)$, f possède aussi un minimum local en $(-2,2)$.

En conclusion, f possède deux minimums locaux sur $\mathbb{R}^2$; ils sont obtenus en $(-2,2)$ et $(2,-2)$ et, en ces points, f vaut -32.

2) *Nous allons montrer que ce minimum est en fait le minimum de f sur $\mathbb{R}^2$ tout entier.*
En effet, il est clair que, si $\|(x,y)\|_2 \to +\infty$, alors $f(x,y) \to +\infty$ (en posant $x = r\cos\theta$ et $y = r\sin\theta$).
Par définition de la limite, on a donc :

$$\exists\, r > 0,\ \|(x,y)\|_2 > r \implies f(x,y) > 0\,.$$

Or la boule $\mathcal{B}_f(0,r)$ est une partie fermée bornée, donc on sait que f admet sur cette boule un minimum, que ce minimum est atteint, et qu'il est nécessairement inférieur à $f((0,0) = 0$. Ce sera donc forcément un minimum global de f sur $\mathbb{R}^2$.
Un minimum global étant aussi local, ce minimum est à chercher parmi les deux trouvés précédemment.
La valeur de f en ces deux points étant la même, f admet en ces deux points un minimum global.

SF18.12 Utiliser la continuité sur un compact

Lorsque f est une application de classe $\mathscr{C}^1$ définie sur une partie compacte K, on utilise le théorème :

> Toute fonction continue sur une partie compacte d'un espace vectoriel normé et à valeurs dans $\mathbb{R}$ est bornée et atteint ses bornes.

On cherchera donc les extrema de f sur K parmi :

- ses extrema sur le bord ∂K ;
- et ses extrema sur l'intérieur $\overset{\circ}{K}$, qui sont à chercher parmi les points critiques de f sur $\overset{\circ}{K}$.

Exemple

Soit f définie sur $\mathbb{R}^2$ par : $\forall\,(x,y) \in \mathbb{R}^2$, $f(x,y) = \dfrac{x+y}{(1+x^2)(1+y^2)}$.

Déterminer le maximum de f sur $K = [0\,;1]^2$.

La fonction f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. K est une partie fermée bornée de $\mathbb{R}^2$ donc une partie compacte. Sur K, f est bornée et atteint ses bornes, elle admet donc en particulier un maximum.

- *Cherchons d'abord les points critiques de f sur $\overset{\circ}{K}$. Il sont donnés par*
$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = 0 \quad \text{et} \quad (x,y) \in]0\,;1[^2$$
soit
$$\begin{cases} (1+x^2) - 2x(x+y) = 0 \\ (1+y^2) - 2y(x+y) = 0 \end{cases} \quad \text{et} \quad (x,y) \in]0\,;1[^2$$
qui conduit à : $\dfrac{2x}{1+x^2} = \dfrac{2y}{1+y^2}$ *d'où* $(x-y)(1-xy) = 0$ *puis* $x = y$ *puisque* $xy \neq 1$ *et enfin :* $x = y = \dfrac{\sqrt{3}}{3}$.
- *On étudie maintenant f sur le bord de K, c'est-à-dire sur les segments $[OA]$, $[AB]$, $[BC]$ et $[OC]$ où $O = (0,0)$, $A = (1,0)$, $B = (1,1)$ et $C = (0,1)$.*
  - *Sur $[OA]$: on a, pour $0 \leqslant x \leqslant 1$: $f(x,0) = \dfrac{x}{1+x^2}$, qui atteint son maximum en $x = 1$, et ce maximum vaut $\dfrac{1}{2}$.*
  - *Sur $[AB]$: on a, pour $0 \leqslant y \leqslant 1$: $f(1,y) = \dfrac{1+y}{2(1+y^2)}$, et l'étude de la fonction $y \mapsto f(1,y)$ montre qu'elle atteint son maximum sur $[0\,;1]$ en $y = \sqrt{2}-1$, ce maximum étant égal à $\dfrac{1+\sqrt{2}}{4}$.*
  - *On obtient des résultats symétriques sur les segments $[BC]$ et $[OC]$.*
- *Finalement, le maximum de f sur K est égal au maximum des trois valeurs :*
$$\frac{1}{2},\ \frac{1+\sqrt{2}}{4} \quad \text{et} \quad f\left(\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3}\right) = 3\frac{\sqrt{3}}{8}.$$
Ce maximum vaut $3\dfrac{\sqrt{3}}{8}$, et il est atteint au point $\left(\dfrac{\sqrt{3}}{3},\dfrac{\sqrt{3}}{3}\right)$.

■ Comment utiliser le calcul différentiel en géométrie ?

SF18.13 Utiliser les formules du cours

Il s'agit essentiellement d'appliquer les résultats rappelés en **[S18.10]** et **[S18.11]**, permettant de trouver l'équation de la tangente à une courbe plane ou l'équation du plan tangent à une surface.

✓ Il ne reste pratiquement plus de géométrie au programme. Mais on vous demande quand même de savoir faire des choses simples, comme trouver et utiliser des équations de droites et de plans.

Exemple

Soit (H) l'hyperbole d'équation $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$.

Donner une condition nécessaire et suffisante portant sur les réels u, v, h pour que la droite d'équation $ux + vy + h = 0$ soit tangente à (H).

L'équation de (H) *peut s'écrire* $f(x,y)=0$ *avec :*

$$f(x,y)=\frac{x^2}{a^2}-\frac{y^2}{b^2}-1\,.$$

f *est de classe* $\mathscr{C}^1$ *et le gradient de* f *en un point* $M_0(x_0,y_0)\in(H)$ *est le vecteur :*

$$\nabla f(x_0,y_0)=\left(\frac{2x_0}{a^2},-\frac{2y_0}{b^2}\right),$$

qui n'est jamais nul puisque le point de coordonnées $(0,0)$ *n'appartient pas à* (H)*. On en déduit que l'équation de la tangente en* M_0 *est :*

$$\frac{2x_0}{a^2}(x-x_0)-\frac{2y_0}{b^2}(y-y_0)=0\,,$$

soit, compte tenu de $\dfrac{x_0^2}{a^2}-\dfrac{y_0^2}{b^2}=1$ *:*

$$\frac{x_0}{a^2}\,x-\frac{y_0}{b^2}\,y=1\,.$$

La droite d'équation $ux+vy+h=0$ *est donc tangente à* (H) *si et seulement si* $h\neq0$ *et s'il existe* $M_0(x_0,y_0)$ *appartenant à* (H) *tel que :*

$$\frac{u}{h}=-\frac{x_0}{a^2}\quad\text{et}\quad\frac{v}{h}=\frac{y_0}{b^2}\,,$$

ce qui équivaut à :

$$h\neq0\quad\text{et}\quad a^2u^2-b^2v^2=h^2\,.$$

Exemple

Soit (S) la surface d'équation cartésienne $z=\dfrac{2xy}{x^2+y^2}$.

Déterminer l'intersection de (S) avec son plan tangent au point $A\left(1,\sqrt{3},\dfrac{\sqrt{3}}{2}\right)$.

- (S) *est la surface d'équation* $f(x,y,z)=0$ *avec* $f(x,y,z)=(x^2+y^2)z-2xy$*.* f *est de classe* $\mathscr{C}^1$ *et son gradient est égal à :*

  $$\nabla f(x,y,z)=\left(2(xz-y),2(yz-x),x^2+y^2\right)\,,$$

  donc

  $$\nabla f(A)=\left(-\sqrt{3},1,4\right).$$

  Ce vecteur étant non nul, l'équation du plan tangent en A *est donc :*

  $$-\sqrt{3}(x-1)+(y-\sqrt{3})+4\left(z-\frac{\sqrt{3}}{2}\right)=0\ \text{ou :}\ \sqrt{3}x-y-4z+2\sqrt{3}=0\,.$$

- *Un point* $M(x,y,z)$ *appartient à l'intersection de ce plan avec* (S) *si et seulement si il vérifie :*

  $$z=\frac{2xy}{x^2+y^2}\quad\text{et}\quad\sqrt{3}x-y-4z+2\sqrt{3}=0\,,$$

  ce qui équivaut à :

  $$\begin{cases}(\sqrt{3}x-y)(x^2+y^2)+2\sqrt{3}(x^2+y^2)-8xy=0\\ \sqrt{3}x-y-4z+2\sqrt{3}=0\end{cases}$$

ou encore :

$$\begin{cases} (\sqrt{3}x - y)(x^2 + y^2 + 2x - 2\sqrt{3}y) = 0 \\ \sqrt{3}x - y - 4z + 2\sqrt{3} = 0 \end{cases},$$

soit finalement :

$$\begin{cases} \sqrt{3}x - y = 0 \\ \sqrt{3}x - y - 4z + 2\sqrt{3} = 0 \end{cases} \quad \textit{ou} \quad \begin{cases} x^2 + y^2 + 2x - 2\sqrt{3}y = 0 \\ \sqrt{3}x - y - 4z + 2\sqrt{3} = 0 \end{cases}.$$

Le premier système correspond à l'intersection de deux plans non parallèles, c'est donc une droite ; le second système correspond à l'intersection du cylindre de base le cercle du plan xOy d'équation $(x+1)^2 + \left(y - \sqrt{3}\right)^2 = 4$ et d'axe parallèle à Oz avec un plan, c'est une ellipse.

À vous de jouer !

Exercice 18.1 **SF18.3** **SF18.4**

Soient a, b, c, d quatre réels strictement positifs et f l'application définie sur $\mathbb{R}^2$ par :

$$f(x,y) = \begin{cases} \dfrac{|x|^a\,|y|^b}{|x|^c + |y|^d} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0\,. \end{cases}$$

Montrer que f est continue sur $\mathbb{R}^2$ si et seulement si $\dfrac{a}{c} + \dfrac{b}{d} > 1$.

Exercice 18.2 **SF18.3** **SF18.4** **SF18.8**

Soit a un nombre réel, et f l'application définie sur $\mathbb{R}^2$ par :

$$f(x,y) = \begin{cases} (x^2+y^2)^a \sin\left(\dfrac{1}{\sqrt{x^2+y^2}}\right) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } x = y = 0\,. \end{cases}$$

1) Montrer que f est continue sur $\mathbb{R}^2$ si et seulement si $a > 0$.

2) Montrer que f est différentiable sur $\mathbb{R}^2$ si et seulement si $a > \dfrac{1}{2}$.

3) Montrer que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ si et seulement si $a > 1$.

Exercice 18.3 **SF18.8**

L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure euclidienne canonique.
Montrer que l'application $f : x \mapsto \dfrac{1}{\|x\|}$ est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n \setminus \{0\}$ et calculer son gradient.

Exercice 18.4 **SF18.10**

1) Montrer que l'application $\varphi : (x,y) \mapsto \left(x, \dfrac{y}{x}\right)$ définit une bijection entre $\mathbb{R}_+^* \times \mathbb{R}$ et un ouvert de $\mathbb{R}^2$ que l'on précisera.
Vérifier que φ et son inverse sont de classe $\mathscr{C}^1$.

2) Déterminer toutes les fonctions f de $\mathbb{R}_+^* \times \mathbb{R}$ dans $\mathbb{R}$, de classe $\mathscr{C}^1$, qui vérifient :

$$(E_1) \qquad x \frac{\partial f}{\partial x}(x,y) + y \frac{\partial f}{\partial y}(x,y) = 0\,.$$

3) Déterminer toutes les fonctions f de $\mathbb{R}_+^* \times \mathbb{R}$ dans $\mathbb{R}$, de classe $\mathscr{C}^2$, qui vérifient :

$$(E_2) \qquad x^2 \frac{\partial^2 f}{\partial x^2} + 2xy \frac{\partial^2 f}{\partial x \partial y} + y^2 \frac{\partial^2 f}{\partial y^2} = 0\,.$$

Exercice 18.5 **SF18.9** **SF18.11**

On munit l'espace vectoriel $\mathbb{R}^n$ du produit scalaire canonique noté $\langle \cdot | \cdot \rangle$.
Soit $f \in \mathscr{L}(\mathbb{R}^n)$, symétrique, de valeurs propres strictement positives.

1) Montrer que : $\forall\, x \neq 0$, $\langle f(x) | x \rangle > 0$.

2) Pour $u \in \mathbb{R}^n$, on définit g par $g(x) = \dfrac{1}{2} \langle f(x) | x \rangle - \langle u | x \rangle$.
Montrer que g est différentiable en tout point de $\mathbb{R}^n$ et expliciter sa différentielle.

3) Montrer que g admet un unique point critique et que ce point est un minimum global de g.

Exercice 18.6 **SF18.11** **SF18.12**

Soit $f(x,y) = \sqrt{x^2+y^2} + x^2$.
Trouver les extrema de f sur $D = \left\{(x,y) \in \mathbb{R}^2,\ x^2 + 2y^2 \leqslant 8\right\}$.

Exercice 18.7 **SF18.13**

Trouver les plans tangents à la surface (S) d'équation $x^2 + y^2 + 2z^2 = 1$ et orthogonaux à la droite d'équation $x = \dfrac{y}{3} = -\dfrac{z}{2}$.

CORRIGÉS DES EXERCICES

Exercice 1.1

1) H étant déjà non vide et stable par la loi $*$, pour montrer qu'il s'agit d'un sous-groupe de G, il ne reste plus à montrer que pour tout $h \in H$, $h^{-1} \in H$.

Soit donc $h \in H$.

L'application $\gamma_h : \begin{cases} H & \longrightarrow & H \\ x & \longmapsto & h*x \end{cases}$ (c'est bien une application de H dans H puisque H stable) est une application injective, car $h*x = h*y \implies x = y$, tout élément d'un groupe étant régulier.

Étant injective de H dans H qui est un ensemble fini, elle est bijective. En particulier, il existe un (et un seul) $x \in H$ tel que $h = \gamma_h(x)$ c'est-à-dire $h = h*x$. Cette égalité implique $x = e_G$ (car h régulier), ce qui prouve que $e_G \in H$.

Ensuite, toujours grâce à la surjectivité de γ_h, il existe un (et un seul) $y \in H$ tel que $\gamma_h(y) = e_G$, c'est-à-dire $e_G = h*y$, soit $y = h^{-1}$. Cela prouve donc que h^{-1} appartient à H, ce qu'il fallait démontrer.

2) Soit H une partie finie non vide de $\mathbb{C}^*$, stable par la multiplication. D'après ce qui précède, il s'agit d'un sous groupe de $(\mathbb{C}^*,\times)$. En particulier son élément neutre est celui de $(\mathbb{C}^*,\times)$, c'est-à-dire 1. Donc si H est d'ordre n, tout élément x de H vérifie $x^n = 1$, donc H est inclus dans le sous-groupe $\mathbb{U}_n$ des racines n^{es} de l'unité.
Et puisque $\operatorname{Card} G = n = \operatorname{Card} \mathbb{U}_n$, on a $H = \mathbb{U}_n$.

Exercice 1.2

Raisonnons par l'absurde et supposons qu'il n'existe pas d'élément d'ordre p.
D'après le théorème de Lagrange, on sait que l'ordre de tout élément divise $2p$, donc est égal à 1 (si c'est l'élément neutre), 2, p ou $2p$.
S'il n'y a pas d'élément d'ordre p, il ne peut pas y en avoir d'ordre $2p$ (car si x est d'ordre $2p$, x^2 est d'ordre p).
Tous les éléments, excepté l'élément neutre, sont donc d'ordre 2 (et $p \geqslant 3$).
Soient x et y deux éléments distincts et différents de e_G. Alors l'élément xy ne peut être égal à x ni à y, et l'on a :

$$(xy)^2 = e_G = x^2y^2\,,$$

d'où $x(yx)y = x(xy)y$ puis $xy = yx$.
Le groupe engendré par x et y est donc égal à $H = \{e_G,x,y,xy\}$. C'est un sous-groupe de G donc d'après le théorème de Lagrange, 4 divise $2p$ donc 2 divise p, ce qui est contradictoire.

On a donc montré par l'absurde que dans un groupe d'ordre $2p$ avec p premier, il existe un élément d'ordre p.

Exercice 1.3

1) a) Pour tout entier k, l'application $f_k : x \mapsto x^k$ est un morphisme du groupe G sur lui-même (en effet, $(xy)^k = x^k y^k$ car G abélien).
G_a (respectivement G_b) est l'image de G par le morphisme f_a (respectivement f_b), c'est donc un sous-groupe de G.

b) Considérons maintenant l'application :

$$\phi : \begin{cases} G_a \times G_b & \longrightarrow & G \\ (y,z) & \longmapsto & yz \end{cases}.$$

ϕ est facilement un morphisme de groupes (là encore, puisque G est abélien).
La question posée équivaut à démontrer que ϕ est bijective, c'est-à-dire que c'est un isomorphisme.

- Soit $(y,z) \in G_a \times G_b$ tel que $\phi(y,z) = e_G$.
  Alors $yz = e_G$ et il existe u et v dans G tels que $y = u^a$ et $z = v^b$.
  On a donc $y^b = u^{ab} = u^n = e_G$ donc l'ordre de y divise b ; de même l'ordre de z divise a. Mais $y = z^{-1}$ donc y et z ont même ordre ; cet ordre étant un diviseur commun à a et à b est égal à 1, ce qui signifie que $y = z = e_G$.
  Ainsi le noyau de ϕ est réduit à $\{(e_G,e_G)\}$ qui est l'élément neutre du groupe produit $G_a \times G_b$, donc ϕ est injective.
- Pour montrer que ϕ est surjective, nous allons utiliser l'identité de Bezout : $a \wedge b = 1$, donc il existe des entiers α et β tels que $\alpha a + \beta b = 1$.
  On peut alors écrire, pour tout $x \in G$:

$$x = \left(x^\alpha\right)^a \left(x^\beta\right)^b = \phi(y,z)\,,$$

  où $y = \left(x^\alpha\right)^a \in G_a$ et $z = \left(x^\beta\right)^b \in G_b$.
  Donc ϕ est surjective de $G_a \times G_b$ sur G, ce qui achève la démonstration.

2) Nous avons déjà indiqué que l'application f_k est bien un morphisme du groupe G dans lui-même.

– Pour montrer que f_k est injective, cherchons son noyau.
Soit $x \in \operatorname{Ker} f_k$; on a donc $x^k = e_G$, mais on a aussi $x^n = e_G$ par hypothèse. L'ordre de x divise donc k et n, et puisque $k \wedge n = 1$, cet ordre est égal à 1 c'est-à-dire que $x = e_G$.
On a donc montré que $\operatorname{Ker} f_k = \{e_G\}$: f_k est injective.

– Puisque $k \wedge n = 1$, il existe des entiers α et β tels que $\alpha k + \beta n = 1$.
On a alors, pour tout $x \in G$:

$$x = x^{\alpha k+\beta n} = \left(x^{\alpha}\right)^k \left(\underbrace{x^n}_{=e_G}\right)^{\beta}$$

$$= \left(x^{\alpha}\right)^k = f_k(x^{\alpha}) = f_k \circ f_{\alpha}(x)\,,$$

donc f_k est bijective de G sur G et $f_k^{-1} = f_{\alpha}$.

Exercice 1.4

1) • Supposons d'abord $n = p^k$ avec p premier et $k \in \mathbb{N}^*$.
Un élément $\overline{x} \in \mathbb{Z}/n\mathbb{Z}$ est inversible si et seulement si $x \wedge n = 1$, c'est-à-dire si $x \wedge p = 1$ ou encore : x non multiple de p (car p premier).
Donc $\overline{x}$ n'est pas inversible si et seulement si $x \in p\mathbb{Z}$; l'ensemble $N(\mathbb{Z}/n\mathbb{Z})$ de ces éléments est donc l'image par la projection canonique $\pi\colon x \mapsto \overline{x}$ de l'idéal $p\mathbb{Z}$ de l'anneau $\mathbb{Z}$.
π étant un morphisme d'anneaux, $N(\mathbb{Z}/n\mathbb{Z})$ est un idéal, c'est-à-dire que l'anneau $\mathbb{Z}/n\mathbb{Z}$ est local.

• Réciproquement, supposons $\mathbb{Z}/n\mathbb{Z}$ local et, par l'absurde, supposons que n admet deux diviseurs premiers distincts p et q. Alors $\overline{p}$ et $\overline{q}$ appartiennent à $N(\mathbb{Z}/n\mathbb{Z})$.
Or p et q sont premiers entre eux donc d'après le théorème de Bezout, il existe des entiers u et v tels que $up + vq = 1$.
On a donc $\overline{u}\,\overline{p} + \overline{v}\,\overline{q} = \overline{1}$ et, par définition d'un idéal, $\overline{1}$ appartient à $N(\mathbb{Z}/n\mathbb{Z})$, ce qui est absurde puisque $\overline{1}$ est inversible.

On a donc montré par l'absurde que n admet un unique facteur premier ; il est donc de la forme $n = p^k$ avec p premier et $k \in \mathbb{N}^*$.

2) • On vérifie facilement que $\mathbb{Z}_p$ est un sous-anneau de $(\mathbb{Q}, +, \times)$.

– $\mathbb{Z}_p$ est inclus dans $\mathbb{Q}$ par définition, et 1 appartient à $\mathbb{Z}_p$.

– Soient x, $y \in \mathbb{Z}_p$. Alors $x = \dfrac{a}{b}$ et $y = \dfrac{a'}{b'}$ avec b, b' premiers avec p. $x+y$ et xy sont alors égaux à une fraction de la forme $\dfrac{n}{bb'}$, et puisque bb' est encore premier avec p, ils appartiennent à $\mathbb{Z}_p$.

• Un élément $\dfrac{a}{b}$ est inversible dans $\mathbb{Z}_p$ si et seulement si il est non nul et si $\dfrac{b}{a}$ appartient à $\mathbb{Z}_p$, c'est-à-dire si a est premier avec p.
L'ensemble $N(\mathbb{Z}_p)$ des éléments de $\mathbb{Z}_p$ non inversibles dans $\mathbb{Z}_p$ est donc l'ensemble des $\dfrac{a}{b}$ avec $b \wedge p = 1$ et a multiple de p (car p premier, donc a non premier avec p équivaut à a multiple de p). Il s'agit donc de l'ensemble $p\mathbb{Z}_p$, c'est-à-dire l'idéal principal engendré par p.
$N(\mathbb{Z}_p)$ est donc un idéal, $\mathbb{Z}_p$ est un anneau local.

Exercice 1.5

• On commence par résoudre le système :

$$x \equiv 2 \pmod 7 \ ; \ x \equiv 1 \pmod 8\,.$$

9 est une solution évidente, donc d'après le théorème chinois (7 et 8 étant premiers entre eux), l'ensemble des solutions de ce premier système est :

$$\{9 + 56k \mid k \in \mathbb{Z}\}\,.$$

• Pour résoudre le système proposé, il reste à chercher les nombres de la forme $9 + 56k$ qui sont congrus à 3 modulo 9, ce qui revient à chercher les $k \in \mathbb{Z}$ tels que $6 + 56k \equiv 0 \pmod 9$.
Dans $\mathbb{Z}/9\mathbb{Z}$, cela équivaut à :

$$\overline{6 + 56k} = \overline{0} \text{ soit } \overline{2}\,\overline{k} = -\overline{6}\,.$$

$2 \wedge 9 = 1$ donc $\overline{2}$ est inversible dans $\mathbb{Z}/9\mathbb{Z}$; l'équation précédente a donc une solution unique, qui est évidemment $\overline{k} = -\overline{3} = \overline{6}$.
Il existe donc $p \in \mathbb{Z}$ tel que $k = 6 + 9p$; alors $9 + 56k = 9 + 56(6 + 9p)$. L'ensemble des solutions est donc l'ensemble :

$$\{345 + 504p \mid p \in \mathbb{Z}\}\,.$$

Exercice 1.6

On se place dans $\mathbb{Z}/7\mathbb{Z}$, et on y calcule les cubes des éléments :

$$\overline{0}^3 = \overline{0}\ ;\ \overline{1}^3 = \overline{1}\ ;\ \overline{2}^3 = \overline{1}\ ;\ \overline{3}^3 = -\overline{1}\ ;$$
$$\overline{4}^3 = \overline{1}\ ;\ \overline{5}^3 = -\overline{1}\ ;\ \overline{6}^3 = -\overline{1}\,.$$

Les cubes possibles sont donc $\overline{0}$, $\overline{1}$ et $-\overline{1}$. Si $a^3+b^3+c^3$ est multiple de 7 alors dans $\mathbb{Z}/7\mathbb{Z}$ on a $\overline{a}^3+\overline{b}^3+\overline{c}^3=\overline{0}$, et d'après les calculs précédents, l'un des 3 éléments $\overline{a}$, $\overline{b}$ ou $\overline{c}$ est forcément égal à $\overline{0}$.
L'un des trois nombres a, b ou c est donc multiple de 7, et par suite, le produit abc aussi.

Exercice 1.7

- 13 étant premier, on a $a^{13}-a$ divisible par 13 d'après le petit théorème de Fermat.
- De même, 7 étant premier, a^7-a est divisible par 7, donc $a^{13}-a=(a^7-a)(a^6+1)$ aussi.
- De même, 3 étant premier, a^3-a est divisible par 3, donc $a^{13}-a=(a^3-a)(a^{10}+a^8+a^6+a^4+a^2+1)$ aussi.
- Enfin, $a^{13}-a$ est évidemment pair.

$a^{13}-a$ étant divisible par 2, 3, 7 et 13, qui sont premiers entre eux deux à deux, l'est par leur produit, c'est-à-dire par 546.

Exercice 1.8

Procédons par récurrence sur n et notons $x_n = 2^{2^{6n+2}}+3$.
• Pour $n=0$, $x_0=19$ donc la propriété est vérifiée.

• Supposons démontré $x_{n-1}\equiv 0 \pmod{19}$, c'est-à-dire $2^{2^{6n-4}}\equiv -3 \pmod{19}$.
Alors

$$\begin{aligned} 2^{2^{6n+2}} &= \left(2^{2^{6n-4}}\right)^{2^6} \\ &\equiv (-3)^{64} \pmod{19}\,. \end{aligned}$$

D'après le théorème de Fermat, $(-3)^{18}\equiv 1 \pmod{19}$ donc

$$(-3)^{64} = (-3)^{18\times 3+10} \equiv 3^{10} \pmod{19}\,.$$

Or

$$3^3 = 27 \equiv 8 \pmod{19}$$

puis

$$8^2 = 64 \equiv 7 \pmod{19}$$

puis

$$8^3 \equiv 56 \equiv -1 \pmod{19}\,.$$

Finalement

$$3^{10} = 3\left(\left(3^3\right)^3\right) \equiv -3 \pmod{19}\,,$$

et

$$2^{2^{6n+2}} \equiv -3 \pmod{19}\,,$$

ce qui achève la récurrence.

Exercice 1.9

1) Le résultat étant immédiat si $p=2$ ou $p=3$, nous pouvons supposer $p\geqslant 5$.
p étant premier, $\mathbb{Z}/p\mathbb{Z}$ est un corps ; on a donc, pour $x\in(\mathbb{Z}/p\mathbb{Z})^*$:

$$\begin{aligned} x^2=\overline{1} &\iff (x-\overline{1})(x+\overline{1})=\overline{0} \\ &\iff x\in\left\{-\overline{1},\overline{1}\right\}\,. \end{aligned}$$

Les seuls éléments égaux à leurs inverses ($x=x^{-1}$) dans $(\mathbb{Z}/p\mathbb{Z})^*$ sont donc $\overline{1}$ et $-\overline{1}=\overline{p-1}$. On peut donc regrouper les autres éléments (il en existe, car $p\geqslant 5$) deux par deux, d'où l'on tire, en faisant le produit de tous les éléments de $(\mathbb{Z}/p\mathbb{Z})^*$:

$$\overline{(p-1)!} = \overline{1}.(-\overline{1}) \prod_{x\neq\pm 1}\left(\overline{x}\,\overline{x}^{-1}\right) = -\overline{1}\,,$$

soit $(p-1)!\equiv -1 \pmod{p}$.

2) Supposons $(n-1)!\equiv -1 \pmod{n}$; alors dans l'anneau $(\mathbb{Z}/n\mathbb{Z},+,\times)$, on a :

$$\overline{1}\times\overline{2}\times\cdots\times\overline{n-1} = -\overline{1}\,,$$

donc tous les éléments non nuls de $\mathbb{Z}/n\mathbb{Z}$ sont inversibles. $\mathbb{Z}/n\mathbb{Z}$ est donc un corps, c'est-à-dire que n est premier.

Exercice 1.10

1) • A et B étant premiers entre eux, le théorème de Bezout donne l'existence de deux polynômes U et V tels que

$$AU+BV=1 \quad (*)$$

La division euclidienne de U par B s'écrit :

$$U = BQ_1+U_0 \text{ avec } \deg U_0 < \deg B\,,$$

et celle de V par A s'écrit :

$$V = AQ_2+V_0 \text{ avec } \deg V_0 < \deg A\,.$$

En remplaçant dans $(*)$ on obtient :

$$AU_0 + BV_0 + AB(Q_1 + Q_2) = 1\,.$$

Or $\deg(AU_0 + BV_0) < \deg(AB)$ donc cette relation implique $Q_1 + Q_2 = 0$ (sinon le degré de $AB(Q_1 + Q_2)$ serait supérieur ou égal au degré de AB) donc on a :

$$AU_0 + BV_0 = 1$$

avec U_0, V_0 satisfaisant aux conditions demandées.

• Démontrons maintenant l'unicité de (U_0,V_0). S'il existe un autre couple (U_1,V_1) vérifiant les mêmes propriétés, on a :

$$A(U_0 - U_1) = B(V_1 - V_0)\,,$$

et puisque $A \wedge B = 1$, le théorème de Gauss donne $B \big| (U_0 - U_1)$. Mais $\deg(U_0 - U_1) < \deg B$ donc $U_0 - U_1 = 0$ d'où $U_1 = U_0$ puis $V_1 = V_0$.

• Si (U,V) est un couple quelconque de polynômes tels que $AU + BV = 1$ alors :

$$A(U - U_0) = B(V_0 - V)\,.$$

et comme ci-dessus, B divise $U - U_0$. Il existe donc un polynôme P tel que $U = U_0 + BP$, d'où en remplaçant, $V = V_0 + AP$.
Les couples solutions sont donc de la forme :

$$(U,V) = (U_0 + BP,V_0 + AP)\,,$$

avec P polynôme quelconque.

2) Il suffit de développer la relation :

$$\big((1 - X) + X\big)^{2n-1} = 1\,,$$

par la formule du binôme :

$$1 = \sum_{k=0}^{2n-1} \tbinom{2n-1}{k}(1 - X)^k X^{2n-1-k}$$

$$= X^n \underbrace{\left(\sum_{k=0}^{n-1} \tbinom{2n-1}{k} X^{n-1-k}(1-X)^k\right)}_{=U_0}$$

$$+(1-X)^n \underbrace{\left(\sum_{k=n}^{2n-1} \tbinom{2n-1}{k} X^{2n-1-k}(1-X)^{k-n}\right)}_{=V_0}$$

Donc $U_0 = \sum_{k=0}^{n-1} \tbinom{2n-1}{k} X^{n-1-k}(1 - X)^k$ et après le changement d'indice $k' = 2n-1-k$, on constate que $V_0(X) = U_0(1-X)$ (ce qui n'a rien de surprenant !).

Exercice 1.11

• Parmi les polynômes constants, seuls $P = 0$ et $P = 1$ conviennent.

• Supposons désormais P non constant, et soit $\alpha \in \mathbb{C}$ une racine de P. Alors α^2 est aussi racine de P, puis aussi α^4 etc...
Ainsi, tous les complexes α^{2^n} $(n \in \mathbb{N}^*)$ sont racines de P. Le nombre de racines de P étant fini, on en déduit facilement que $\alpha = 0$ ou que α est une racine de l'unité (en effet, il existe $n > m$ tels que $\alpha^{2^n} = \alpha^{2^m}$ donc $\alpha = 0$ ou $\alpha^{2^n - 2^m} = 1$), et en particulier, $\alpha = 0$ ou $|\alpha| = 1$.

D'autre part, on a aussi $P\big((\alpha + 1)^2\big) = 0$, donc $(\alpha + 1)^2$ est aussi racine de P, donc d'après ce qui précède, $\alpha = -1$ ou $|\alpha + 1| = 1$.
Finalement les racines possibles de P sont 0, -1, et les solutions de $|\alpha| = |\alpha + 1| = 1$, c'est-à-dire j et j^2 (cela se justifie soit par le calcul, soit géométriquement (intersection de deux cercles)).
On a donc :

$$P = \lambda X^a (X + 1)^b (X - \mathrm{j})^c (X - \mathrm{j}^2)^d\,.$$

Il ne reste plus qu'à remplacer dans l'équation initiale et on trouve finalement :

$$P = (X - \mathrm{j})^c (X - \mathrm{j}^2)^c = (X^2 + X + 1)^c\,,$$

avec $c \in \mathbb{N}$.

Exercice 1.12

1) Si A possède n racines distinctes $x_1 < x_2 < \ldots < x_n$, le théorème de Rolle appliqué à P sur chaque intervalle $[x_i\,;x_{i+1}]$ assure l'existence de $n - 1$ racines réelles distinctes de P'. Puisque $\deg P' = n - 1$, il est scindé dans $\mathbb{R}[X]$.

2) Avec les notations précédentes, on a $P = a_n \prod_{i=1}^{n} (X - x_i)$ d'où, en dérivant ce produit :

$$\frac{P'}{P} = \sum_{i=1}^{n} \frac{1}{X - x_i}\,.$$

En dérivant à nouveau :

$$\frac{PP'' - P'^2}{P^2} = -\sum_{i=1}^{n} \frac{1}{(X - x_i)^2}\,.$$

Pour $x \in \mathbb{R}$, distinguons alors deux cas :

– s'il existe i tel que $x = x_i$, puisque x_i n'est pas racine de P', on a :

$$P'^2(x) - P(x)P''(x) = P'^2(x_i) > 0\,;$$

– sinon, en utilisant le calcul précédent :

$$(P'^2 - PP'')(x) = P^2(x)\sum_{i=1}^{n}\frac{1}{(x-x_i)^2} > 0.$$

Dans les deux cas, on a bien le résultat demandé.

3) En appliquant le résultat précédent aux polynômes dérivés successifs de P, ce qui est possible car ils sont tous scindés à racines simples d'après la première question, on obtient, pour tout $k \in [\![1\,;n-1]\!]$ et tout $x \in \mathbb{R}$:

$$\left(P^{(k)}(x)\right)^2 - P^{(k-1)}(x)P^{(k+1)}(x) > 0\,.$$

Or $a_k = \dfrac{P^{(k)}(0)}{k!}$ d'après la formule de Taylor, donc, en prenant $x = 0$ dans l'inégalité précédente :

$$(k!)^2 a_k^2 - (k-1)!(k+1)!a_{k-1}a_{k+1} > 0\,,$$

ce qui donne le résultat voulu.

Exercice 1.13

$$(x+1)^n = \mathrm{e}^{2\mathrm{i}n\alpha} \iff x+1 = \mathrm{e}^{2\mathrm{i}\alpha}\mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}\,,$$

avec $k \in [\![0\,;n-1]\!]$.
Les racines sont donc les complexes :

$$\begin{aligned} x_k &= \mathrm{e}^{2\mathrm{i}(\alpha+\frac{k\pi}{n})} - 1 \\ &= 2\mathrm{i}\mathrm{e}^{\mathrm{i}(\alpha+\frac{k\pi}{n})}\sin\left(\alpha+\frac{k\pi}{n}\right). \end{aligned}$$

Le produit des racines est donc :

$$\begin{aligned} \prod_{k=0}^{n-1} x_k &= \prod_{k=0}^{n-1} 2\mathrm{e}^{\mathrm{i}(\alpha+\frac{k\pi}{n}+\frac{\pi}{2})}\sin\left(\alpha+\frac{k\pi}{n}\right) \\ &= 2^n\mathrm{e}^{\mathrm{i}n\alpha}\mathrm{e}^{\mathrm{i}\frac{n(n-1)\pi}{2n}}\mathrm{e}^{\mathrm{i}\frac{n\pi}{2}}\prod_{k=0}^{n-1}\sin\left(\alpha+\frac{k\pi}{n}\right) \\ &= 2^n\mathrm{e}^{\mathrm{i}n\alpha}(-1)^n(-\mathrm{i})\prod_{k=0}^{n-1}\sin\left(\alpha+\frac{k\pi}{n}\right), \end{aligned}$$

et d'autre part, d'après les relations coefficients-racines, ce produit est égal à

$$(-1)^n\left(1-\mathrm{e}^{2\mathrm{i}n\alpha}\right) = (-1)^{n-1}2\mathrm{i}\mathrm{e}^{\mathrm{i}n\alpha}\sin(n\alpha),$$

d'où la formule :

$$\prod_{k=0}^{n-1}\sin\left(\alpha+\frac{k\pi}{n}\right) = \frac{\sin n\alpha}{2^{n-1}}$$

(formule qu'il est facile de vérifier pour $n=1$ et $n=2$).

Exercice 1.14

• Démontrons déjà que le polynôme $P = X^4 - 5X^2 + 1$ n'a pas de racine dans $\mathbb{Q}$.
En effet, si $\dfrac{p}{q}$ était racine de P, avec $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$ et $p \wedge q = 1$, on aurait :

$$p^4 - 5p^2q^2 + q^4 = 0\,,$$

d'où q divise p^4 et p divise q^4. Or $p \wedge q = 1$ donc la seule solution rationnelle possible est ± 1, et puisque ± 1 n'est pas racine, P n'a pas de racine rationnelle.
• Il en résulte que P ne peut pas avoir de facteur irréductible de la forme $X - r$ avec $r \in \mathbb{Q}$. Si P n'était pas irréductible dans $\mathbb{Q}[X]$ il s'écrirait donc $P = AB$ avec A, $B \in \mathbb{Q}[X]$ unitaires de degré 2.
Il existerait donc $(a,b,c,d) \in \mathbb{Q}^4$ tel que :

$$P = (X^2 + aX + b)(X^2 + cX + d)\,,$$

ce qui conduit au système :

$$\left\{\begin{array}{rcl} a+c & = & 0 \\ ac+b+d & = & -5 \\ ad+bc & = & 0 \\ bd & = & 1 \end{array}\right.,$$

soit :

$$\left\{\begin{array}{rcl} c & = & -a \\ -a^2+b+d & = & -5 \\ a(d-b) & = & 0 \\ bd & = & 1 \end{array}\right.\,.$$

– Le cas $a = 0$ implique :

$$b + d = -5\ ;\ bd = 1\,,$$

donc b et d racines du polynôme $X^2 + 5X + 1$; ce polynôme n'ayant pas de racine rationnelle, c'est impossible.

– Le cas $b = d$ conduit à $b = \pm 1$ puis $a^2 = 2b + 5 \in \{3,7\}$. $\sqrt{3}$ et $\sqrt{7}$ étant irrationnels, ce cas aussi est impossible.

Nous avons donc montré que P est irréductible dans $\mathbb{Q}[X]$.

Exercice 2.1

1) Les applications $f \mapsto f(0)$, $f \mapsto f'(0)$ et $f \mapsto \displaystyle\int_0^1 f$ sont des formes linéaires sur E.

Leurs noyaux :

$$H = \{f \in E \mid f(0) = 0\},$$
$$H' = \{f \in E \mid f'(0) = 0\}$$
$$\text{et } H'' = \left\{f \in E \;\middle|\; \int_0^1 f = 0\right\}$$

sont donc des sous-espaces vectoriels de E (ce sont des hyperplans).
Puisque $F = H \cap H' \cap H''$, c'est aussi un sous-espace vectoriel de E.

2) Notons $\mathbb{R}_2[x]$ l'ensemble des fonctions polynômes de degré $\leqslant 2$ définies sur $[0\,;1]$, et montrons que c'est un supplémentaire de F dans E.

• C'est bien un sous-espace vectoriel de E : c'est celui engendré par les fonctions $x \mapsto 1$, $x \mapsto x$ et $x \mapsto x^2$, qui sont bien de classe $\mathscr{C}^1$ sur $[0\,;1]$.

• Soit $f \in F \cap \mathbb{R}_2[x]$.
Alors il existe des réels a, b et c tels que $f(x) = ax^2 + bx + c$ pour tout $x \in [0\,;1]$; les relations $f(0) = f'(0) = 0$ donnent alors $a = b = 0$, puis la relation $\int_0^1 f = 0$ donne ensuite $c = 0$.

Ainsi, $F \cap \mathbb{R}_2[x] = \{0\}$: la somme $F + \mathbb{R}_2[x]$ est directe.

• Il reste à prouver que $F + \mathbb{R}_2[x] = E$, c'est-à-dire que, étant donné $f \in E$, on peut trouver $g \in F$ et $(a,b,c) \in \mathbb{R}^3$ tels que :

$$\forall\, x \in [0\,;1],\ f(x) = g(x) + ax^2 + bx + c\,.$$

Les conditions $g(0) = 0$ et $g'(0) = 0$ donnent $c = f(0)$ et $b = f'(0)$; puis la condition $\int_0^1 g = 0$ donne :

$$\frac{a}{3} + \frac{b}{2} + c = \int_0^1 f\,,$$

d'où $a = 3\left(\int_0^1 f - \frac{1}{2}f'(0) - f(0)\right)$.

Réciproquement, si l'on considère $h : x \mapsto ax^2 + bx + c$ où a, b, c sont donnés par les formules précédentes puis $g = f - h$, le même calcul que ci-dessus montre que l'on aura bien $g \in F$.

On peut donc écrire toute application $f \in E$ sous la forme $f = g + h$ avec $g \in F$ et $h \in \mathbb{R}_2[x]$.

• Conclusion : un supplémentaire de F dans E est $\mathbb{R}_2[x]$.

Exercice 2.2

Nous donnons deux démonstrations du résultat demandé.

1) *Par récurrence sur n*

• Pour $n = 1$ le résultat est immédiat car la famille est alors formée d'un seul polynôme non nul.

• Supposons le résultat démontré à l'ordre $n-1$ ($n \geqslant 2$), et soient $a_1, \ldots, a_n$ n réels distincts. Soient $\lambda_1, \ldots, \lambda_n$ des réels tels que :

$$\sum_{i=1}^{n} \lambda_i (X - a_i)^n = 0 \quad (1)\,.$$

En dérivant cette relation on obtient :

$$\sum_{i=1}^{n} \lambda_i (X - a_i)^{n-1} = 0 \quad (2)\,,$$

et en calculant $(1) - (X - a_n) \times (2)$, ce qui permet d'éliminer le terme $(X - a_n)^n$, on trouve :

$$\sum_{i=1}^{n-1} \lambda_i [(X - a_i)^n - (X - a_n)(X - a_i)^{n-1}] = 0$$

soit : $\sum_{i=1}^{n-1} \lambda_i (a_n - a_i)(X - a_i)^{n-1} = 0\,.$

La propriété au rang $n-1$ implique alors $\lambda_i(a_n - a_i) = 0$ pour tout $i \in [\![1\,;n-1]\!]$ d'où $\lambda_i = 0$ puisque $a_n \neq a_i$.

Enfin, en remplaçant dans (1) on obtient $\lambda_n = 0$. Finalement, tous les coefficients λ_i sont nuls, ce qui montre que la famille $\big((X - a_i)^n\big)_{1 \leqslant i \leqslant n}$ est libre, et achève la récurrence.

2) *À l'aide d'un déterminant*

D'après la formule du binôme, pour tout $j \in [\![1\,;n]\!]$:

$$(X - a_j)^n = \sum_{i=0}^{n} \binom{n}{i} (-a_j)^i X^{n-i}$$

donc la matrice de la famille des vecteurs $(X - a_j)^n$ dans la base $\big(X^n, X^{n-1}, \ldots, X, 1\big)$ de $\mathbb{R}_n[X]$, qui est une matrice de type $(n+1, n)$, s'écrit :

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ & \cdots & \\ \binom{n}{i}(-a_1)^i & \cdots & \binom{n}{i}(-a_n)^i \\ & \cdots & \\ (-a_1)^n & \cdots & (-a_n)^n \end{pmatrix}.$$

Le déterminant extrait formé des n premières lignes est égal à :

$$\begin{vmatrix} 1 & \dots & 1 \\ & \dots & \\ \binom{n}{i}(-a_1)^i & \dots & \binom{n}{i}(-a_n)^i \\ & \dots & \\ n(-a_1)^{n-1} & \dots & n(-a_n)^{n-1} \end{vmatrix}$$

donc à $\left(\prod_{i=1}^{n}\binom{n}{i}\right) V(-a_1,\dots,-a_n)$ où V désigne le déterminant de Vandermonde. Les a_i étant distincts, ce déterminant est non nul ; il en résulte que la matrice extraite de A formée des n premières lignes est inversible, donc A est de rang n.

La famille des n polynômes $(X-a_i)^n$ étant de rang n, elle est libre.

Exercice 2.3

1) Comme indiqué par l'énoncé, procédons par récurrence (finie) sur $p = n - \dim F$.
• Si $p = 0$, $F = G = E$ et un supplémentaire commun à F et G est $\{0\}$.

• Supposons la propriété démontrée à l'ordre $p-1 \geqslant 0$, et soient F et G deux sous-espaces vectoriels de même dimension $n - p < n$.

D'après un résultat rappelé en **[S2.4]**, $F \cup G$ ne peut être égal à E (car $F \cup G = E \implies F = E$ ou $G = E$ ce qui est exclu ici).

Il existe donc $x \in E$ tel que $x \notin F$ et $x \notin G$. Les sommes $F + \mathbb{K}x$ et $G + \mathbb{K}x$ sont donc directes ; si l'on pose $F' = F \oplus \mathbb{K}x$ et $G' = G \oplus \mathbb{K}x$, alors F' et G' sont des sous-espaces vectoriels de E de dimension $n - p + 1 = n - (p-1)$.

D'après le résultat à l'ordre $p-1$, il existe un sous-espace vectoriel H' de E tel que $E = F' \oplus H' = G' \oplus H'$, soit $E = F \oplus \mathbb{K}x \oplus H' = G \oplus \mathbb{K}x \oplus H'$.

En posant $H = H' \oplus \mathbb{K}x$, on a donc bien $E = F \oplus H = G \oplus H$, ce qui achève la récurrence.

2) a) Puisque $E = F \oplus H$, on peut considérer la projection p sur F parallèlement à H. Alors $F = \operatorname{Im} p$ et $H = \operatorname{Ker} p$.

Puisque G est un supplémentaire de $\operatorname{Ker} p$, le théorème d'isomorphisme permet d'affirmer directement que la restriction de p à G est un isomorphisme de G sur $\operatorname{Im} p = F$, ce qui répond à la question.

b) • Il est facile de vérifier que l'application $P \mapsto X^2 P$ est linéaire ; elle réalise alors un isomorphisme de F sur G, puisqu'elle transforme la base $(X^{2k})_{k\in\mathbb{N}}$ de F en la base $(X^{2k})_{k\in\mathbb{N}^*}$ de G.

• Si F et G possédaient un supplémentaire commun H, on pourrait écrire la décomposition du polynôme constant égal à 1 :

$$\exists\,(P_G,P_H) \in G \times H \mid 1 = P_G + P_H\,.$$

Mais $1 \in F$ et $G \subset F$ donc $1 - P_G \in F$; on aurait donc $1 - P_G = P_H \in F \cap H$ d'où $1 - P_G = 0$ puis $1 \in G$, ce qui est faux (les polynômes de G étant tous divisibles par X^2).

On a donc démontré par l'absurde que F et G n'ont pas de supplémentaire commun.

Exercice 2.4

1) On a l'inclusion (importante) :

$$\operatorname{Im}(u+v) \subset \operatorname{Im} u + \operatorname{Im} v\,.$$

En effet, si $y \in \operatorname{Im}(u+v)$, il existe $x \in E$ tel que $y = (u+v)(x) = u(x) + v(x)$ donc $y \in \operatorname{Im} u + \operatorname{Im} v$.
D'où, d'après la formule de Grassmann :

$$\begin{aligned} \operatorname{rg}(u+v) &\leqslant \dim(\operatorname{Im} u + \operatorname{Im} v) \\ &\leqslant \dim \operatorname{Im} u + \dim \operatorname{Im} v \\ &\qquad - \dim(\operatorname{Im} u \cap \operatorname{Im} v) \\ &\leqslant \operatorname{rg} u + \operatorname{rg} v\,. \end{aligned}$$

2) • Si l'on a $\operatorname{rg}(u+v) = \operatorname{rg} u + \operatorname{rg} v$ alors toutes les inégalités précédentes sont des égalités ; on en déduit en particulier $\dim(\operatorname{Im} u \cap \operatorname{Im} v) = 0$ donc $\operatorname{Im} u \cap \operatorname{Im} v = \{0\}$.

Montrons que $\operatorname{Ker}(u+v) = \operatorname{Ker} u \cap \operatorname{Ker} v$:

– l'inclusion $\operatorname{Ker} u \cap \operatorname{Ker} v \subset \operatorname{Ker}(u+v)$ est immédiate ;

– soit $x \in \operatorname{Ker}(u+v)$; alors $u(x) = -v(x)$ donc $u(x) \in \operatorname{Im} u \cap \operatorname{Im} v$ et par suite $u(x) = v(x) = 0$, soit $x \in \operatorname{Ker} u \cap \operatorname{Ker} v$.

On a alors, en utilisant la formule de Grassmann et le théorème du rang :

$$\begin{aligned}\dim(\operatorname{Ker} u + \operatorname{Ker} v) &= \dim(\operatorname{Ker} u) + \dim(\operatorname{Ker} v) \\ &\quad - \dim(\operatorname{Ker} u \cap \operatorname{Ker} v) \\ &= \big(n - \operatorname{rg} u\big) + \big(n - \operatorname{rg} v\big) \\ &\quad - \dim(\operatorname{Ker}(u+v)) \\ &= \big(n - \operatorname{rg} u\big) + \big(n - \operatorname{rg} v\big) \\ &\quad - \big(n - \operatorname{rg}(u+v)\big) \\ &= n = \dim E\,,\end{aligned}$$

ce qui prouve $E = \operatorname{Ker} u + \operatorname{Ker} v$.

• Réciproquement, supposons $\operatorname{Im} u \cap \operatorname{Im} v = \{0\}$ et $E = \operatorname{Ker} u + \operatorname{Ker} v$.

On démontre comme ci-dessus que $\operatorname{Ker}(u+v) = \operatorname{Ker} u \cap \operatorname{Ker} v$ (car la seule relation utilisée était $\operatorname{Im} u \cap \operatorname{Im} v = \{0\}$), d'où :

$$\begin{aligned}\operatorname{rg}(u+v) &= \dim \operatorname{Im}(u+v) \\ &= n - \dim \operatorname{Ker}(u+v) \\ &= n - \dim(\operatorname{Ker} u \cap \operatorname{Ker} v)) \\ &= n - \Big(\dim \operatorname{Ker} u + \dim \operatorname{Ker} v \\ &\qquad - \underbrace{\dim\big(\operatorname{Ker} u + \operatorname{Ker} v\big)}_{=n}\Big) \\ &= 2n - \dim \operatorname{Ker} u - \dim \operatorname{Ker} v \\ &= \operatorname{rg} u + \operatorname{rg} v\,,\end{aligned}$$

ce qui achève de prouver l'équivalence.

3) • L'inclusion $\operatorname{Im}(u \circ v) \subset \operatorname{Im} u$ prouve déjà l'inégalité $\operatorname{rg}(u \circ v) \leqslant \operatorname{rg} u$.

On a aussi $\operatorname{Im}(u \circ v) = u\big(\operatorname{Im} v\big)$ donc $\operatorname{rg}(u \circ v) \leqslant \operatorname{rg} v$ puisque l'image d'un sous-espace vectoriel par une application linéaire a forcément une dimension inférieure.

Finalement on a bien l'inégalité :

$$\operatorname{rg}(u \circ v) \leqslant \min(\operatorname{rg} u, \operatorname{rg} v)$$

(c'est d'ailleurs un résultat du cours).

• On a :

$$\operatorname{rg}(u \circ v) = \dim \operatorname{Im}(u \circ v) = \dim u\big(\operatorname{Im} v\big)\,.$$

En appliquant le théorème du rang à la restriction u' de u à l'image de v, on obtient :

$$\dim \operatorname{Im} v = \dim \operatorname{Ker} u' + \dim \operatorname{Im} u'\,,$$

et puisque $\operatorname{Im} u' = u\big(\operatorname{Im} v\big)$ et que $\operatorname{Ker} u' = \operatorname{Ker} u \cap \operatorname{Im} v$, on en déduit :

$$\begin{aligned}\operatorname{rg} v &= \dim(\operatorname{Ker} u \cap \operatorname{Im} v) + \operatorname{rg}(u \circ v) \\ &\leqslant \dim \operatorname{Ker} u + \operatorname{rg}(u \circ v) \\ &\leqslant n - \operatorname{rg} u + \operatorname{rg}(u \circ v)\,,\end{aligned}$$

ce qui est l'inégalité cherchée.

Exercice 2.5

1) • Supposons $\operatorname{Ker}(g \circ f) = \operatorname{Ker} f$.

Soit $y \in \operatorname{Ker} g \cap \operatorname{Im} f$. Alors :

$$\exists\, x \in E,\ y = f(x \ \text{ et } \ g(y) = 0\,,$$

d'où $g \circ f(x) = 0$, soit $x \in \operatorname{Ker}(g \circ f)$ d'où $x \in \operatorname{Ker} f$ et $y = 0$.

Ainsi $\operatorname{Ker} g \cap \operatorname{Im} f = \{0\}$.

• Supposons $\operatorname{Ker} g \cap \operatorname{Im} f = \{0\}$.

On sait déjà que $\operatorname{Ker} f \subset \operatorname{Ker}(g \circ f)$.

Si $x \in \operatorname{Ker}(g \circ f)$, alors $f(x) \in \operatorname{Ker} g$ donc $f(x) \in \operatorname{Ker} g \cap \operatorname{Im} f$ soit $f(x) = 0$ et $x \in \operatorname{Ker} f$, ce qui donne l'autre inclusion et finalement l'égalité cherchée.

2) • Supposons $\operatorname{Im}(g \circ f) = \operatorname{Im} g$.

Soit $y \in F$. On a $g(y) \in \operatorname{Im} g = \operatorname{Im}(g \circ f)$ donc il existe $x \in E$ tel que $g(y) = g \circ f(x)$.

On a alors : $y = \big(y - f(x)\big) + f(x)$, avec $y - f(x) \in \operatorname{Ker} g$ et $f(x) \in \operatorname{Im} f$, donc $\operatorname{Ker} g + \operatorname{Im} f = F$.

• Supposons $\operatorname{Ker} g + \operatorname{Im} f = F$.

On sait déjà que $\operatorname{Im}(g \circ f) \subset \operatorname{Im} g$.

Soit $z \in \operatorname{Im} g$; il existe $y \in F$ tel que $z = g(y)$; on peut écrire $y = u + v$ avec $u \in \operatorname{Ker} g$ et $v \in \operatorname{Im} f$, et il existe $x \in E$ tel que $v = f(x)$.

Finalement,

$$\begin{aligned}z &= g(u) + g(v) = g(v) \\ &= g \circ f(x) \in \operatorname{Im}(g \circ f),\end{aligned}$$

ce qui donne l'autre inclusion.

Exercice 2.6

Puisque E est de dimension finie, tout sous-espace vectoriel de E admet un supplémentaire. Il existe donc deux sous-espaces vectoriels S et T de E tels que :

$$E = \operatorname{Ker} f \oplus S = \operatorname{Im} f \oplus T\,.$$

• D'après le théorème d'isomorphisme, la restriction f_1 de f à S est un isomorphisme de S sur $\operatorname{Im} f$.

• D'après le théorème du rang :

$$\dim \operatorname{Ker} f = \dim T\,.$$

Les sous-espaces vectoriels $\operatorname{Ker} f$ et T sont donc isomorphes : il existe un isomorphisme u de T sur $\operatorname{Ker} f$.

• On sait alors qu'il existe un endomorphisme g de E tel que :

$$g|_{\operatorname{Im} f} = f_1^{-1} \quad \text{et} \quad g|_T = u$$

(caractérisation d'une application linéaire par sa restriction à deux sous-espaces vectoriels supplémentaires).
Alors :

– g est bijective puisque, par exemple, l'image d'une base de E formée de la réunion de bases de $\operatorname{Im} f$ et de T sera la réunion de deux bases de S et $\operatorname{Ker} f$ respectivement (car f_1^{-1} et u sont des bijections), donc sera une base de E.

– g vérifie bien la relation demandée puisque, si $x \in E$, on peut écrire :

$$x = x_K + x_S \text{ avec } (x_K, x_S) \in \operatorname{Ker} f \times S,$$

puis :

$$f(x) = f(x_K) + f(x_S) = f(x_S) = f_1(x_S),$$

d'où :

$$g \circ f(x) = f_1^{-1} \circ f_1(x_S) = x_S\,,$$

et enfin : $f \circ g \circ f(x) = f(x_S) = f(x)$.

Exercice 2.7

1) Puisque p et q commutent :

$$(p \circ q)^2 = p \circ q \circ p \circ q = p^2 \circ q^2 = p \circ q$$

donc $p \circ q$ est un projecteur par définition.

2) • Soit $x \in \operatorname{Ker} p + \operatorname{Ker} q$. Alors $x = y + z$ avec $y \in \operatorname{Ker} p$ et $z \in \operatorname{Ker} q$.
Donc :

$$\begin{aligned} p \circ q(x) &= p \circ q(y) + p \circ q(z) \\ &= q \circ p(y) + p \circ q(z) = 0\,. \end{aligned}$$

Par suite, $\operatorname{Ker} p + \operatorname{Ker} q \subset \operatorname{Ker}(p \circ q)$.

• Soit $x \in \operatorname{Ker}(p \circ q)$. Alors on peut écrire :

$$x = q(x) + (x - q(x))$$

avec $q(x) \in \operatorname{Ker} p$ et $(x - q(x)) \in \operatorname{Ker} q$.
Donc $\operatorname{Ker}(p \circ q) \subset \operatorname{Ker} p + \operatorname{Ker} q$.

En conclusion : $\operatorname{Ker}(p \circ q) = \operatorname{Ker} p + \operatorname{Ker} q$.

3) • On sait déjà que :

$$\operatorname{Im}(p \circ q) \subset \operatorname{Im} p \quad \text{et} \quad \operatorname{Im}(q \circ p) \subset \operatorname{Im} q\,,$$

donc $\operatorname{Im}(p \circ q) \subset \operatorname{Im} p \cap \operatorname{Im} q$ puisque $p \circ q = q \circ p$.

• Soit $x \in \operatorname{Im} p \cap \operatorname{Im} q$. Alors $p(x) = x$ et $q(x) = x$ donc $p \circ q(x) = x$ et $x \in \operatorname{Im}(p \circ q)$.
Par suite, $\operatorname{Im} p \cap \operatorname{Im} q \subset \operatorname{Im}(p \circ q)$.

En conclusion : $\operatorname{Im}(p \circ q) = \operatorname{Im} p \cap \operatorname{Im} q$.

Exercice 2.8

En utilisant le théorème du rang :

$$\begin{aligned} &\dim \operatorname{Ker} u + \dim \operatorname{Ker}(\mathrm{Id}_E - u) \\ &\quad = n - \operatorname{rg} u + \dim \operatorname{Ker}(v) \\ &\quad = 2n - (\operatorname{rg} u + \operatorname{rg} v) \geqslant n\,. \end{aligned}$$

Mais les sous-espaces vectoriels $\operatorname{Ker} u$ et $\operatorname{Ker}(\mathrm{Id}_E - u)$ sont en somme directe puisque leur intersection est réduite à $\{0\}$ (vérification immédiate).

✓ On peut dire aussi qu'il s'agit de deux sous-espaces propres associés à des valeurs propres distinctes (ici, 0 et 1), donc leur somme est directe.

On a donc forcément :

$$E = \operatorname{Ker} u \oplus \operatorname{Ker}(\mathrm{Id}_E - u)\,.$$

Tout vecteur x de E s'écrit donc de façon unique sous la forme $x = x_1 + x_2$ avec $x_1 \in \operatorname{Ker} u$ et $x_2 \in \operatorname{Ker}(\mathrm{Id}_E - u)$, et on a alors :

$$u(x) = u(x_1) + u(x_2) = x_2\,,$$

c'est-à-dire que u est la projection sur $\operatorname{Ker}(\mathrm{Id}_E - u)$ parallèlement à $\operatorname{Ker} u$ (et v est l'autre projection).

Exercice 2.9

1) • Si $u = 0$ on peut prendre $p = 0$ et tout automorphisme $v \in \mathrm{GL}(E)$ convient.

• Si u est un automorphisme, on peut prendre $p = \mathrm{Id}_E$ et $v = u$.

• Plaçons-nous donc dans le cas où $\operatorname{Ker} u \neq \{0\}$ et $\operatorname{Im} u \neq \{0\}$.
Soit $(e_1, \ldots, e_k)$ une base de $\operatorname{Ker} u$ (avec $1 \leqslant k \leqslant n-1$), que l'on complète en une base $(e_1, \ldots, e_n)$ de E.
D'après le théorème d'isomorphisme, la famille $\big(u(e_{k+1}), \ldots, u(e_n)\big)$ est une base de $\operatorname{Im} u$. Notons $(e'_{k+1}, \ldots, e'_n)$ cette base, que l'on complète en une base $(e'_1, \ldots, e'_n)$ de E.
On considère alors l'endomorphisme v de E défini par :

$$\forall\, i \in [\![1\,;n]\!],\ v(e_i) = e'_i\,.$$

v est un automorphisme, car il transforme une base en une base.

D'autre part, soit p la projection sur $\operatorname{Im} u$ parallèlement à $\operatorname{Vect}(e'_1,\dots,e'_k)$ (cela a bien un sens car ces deux sous-espaces sont supplémentaires).
Alors, pour tout $i \in [\![1\,;k]\!]$:

$$p \circ v(e_i) = p(e'_i) = 0 = u(e_i)\,,$$

et pour tout $i \in [\![k+1\,;n]\!]$:

$$p \circ v(e_i) = p(e'_i) = e'_i = u(e_i)\,.$$

Puisque les endomorphismes u et $p \circ v$ coïncident sur une base, ils sont égaux.

2) • Là encore, les cas $\operatorname{Ker} u = \{0\}$ ou $\operatorname{Im} u = \{0\}$ sont immédiats.

• Sinon, en reprenant les notations précédentes, et en appelant q la projection sur $\operatorname{Vect}(e_{k+1},\dots,e_n)$ parallèlement à $\operatorname{Ker} u$, il est facile de vérifier que l'on a $u = v \circ q$ (en comparant là encore les images des vecteurs de la base $(e_1,\dots,e_n)$).

Exercice 3.1

On note $A = (a_{ij})_{1\leqslant i,j\leqslant n}$.

• On a déjà démontré page 250 la relation :

$$AE_{ij} = \sum_{k=1}^{n} a_{ki} E_{kj}\,,$$

qui montre que $\operatorname{tr}(AE_{ij}) = a_{ji} \quad (*)$.

• Puisque $\operatorname{tr}(M(AN)) = \operatorname{tr}((AN)M)$, la relation de l'énoncé peut s'écrire $\operatorname{tr}(AMN) = \operatorname{tr}(ANM)$.

En appliquant cette relation avec $M = E_{ij}$ et $N = E_{k\ell}$, compte tenu de la formule $E_{ij}E_{kl} = \delta_{jk}E_{i\ell}$, on obtient :

$$\delta_{jk} \operatorname{tr}(AE_{i\ell}) = \delta_{\ell i} \operatorname{tr}(AE_{kj})\,,$$

soit en utilisant $(*)$:

$$\delta_{jk} a_{\ell i} = \delta_{\ell i} a_{jk}\,,$$

cela étant vrai pour tous i, j, k, ℓ de $[\![1\,;n]\!]$. Pour $\ell = i$ on obtient :

$$\forall\,(i,j,k) \in [\![1\,;n]\!]^3,\ a_{jk} = \delta_{jk} a_{ii}$$

donc la matrice A est scalaire.

Exercice 3.2

• Soit p la projection sur H parallèlement à D. Soit $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ et $y = (y_1,\dots,y_n) = p(x)$.

- $p(x) \in H$ donc :

$$u_1y_1 + u_2y_2 + \dots + u_ny_n = 0 \quad (1)$$

- $p(x) - x \in D$ donc il existe un réel λ tel que $p(x) - x = \lambda u$, soit encore :

$$\forall\, i \in [\![1\,;n]\!],\ y_i = x_i + \lambda u_i \quad (2)$$

De (1) et (2) on déduit $\lambda = -\sum_{j=1}^{n} u_j x_j$ puisque $\sum_{j=1}^{n} u_j^2 = 1$. D'où :

$$\forall\, i \in [\![1\,;n]\!],\ y_i = x_i - \sum_{j=1}^{n} u_i u_j x_j\,.$$

La matrice de p dans la base canonique de $\mathbb{R}^n$ est donc la matrice dont le terme d'indice (i,j) est : $\delta_{ij} - u_iu_j$, soit

$$M(p) = \begin{pmatrix} 1-u_1^2 & u_1u_2 & \dots & u_1u_n \\ u_2u_1 & 1-u_2^2 & \dots & u_2u_n \\ \vdots & & & \vdots \\ u_nu_1 & u_nu_2 & \dots & 1-u_n^2 \end{pmatrix}$$

• Si q est la projection sur D parallèlement à H, on a $p + q = \mathrm{Id}_E$. La symétrie par rapport à D de direction H est $s = 2q - \mathrm{Id}_E$, soit $s = \mathrm{Id}_E - 2p$, donc la matrice de s est $\mathrm{I}_n - 2M(p)$.

Exercice 3.3

On calcule $A^2 = \mathrm{I}_n + 2CL + (CL)^2$. Mais LC est une matrice (1,1) c'est-à-dire un scalaire. En notant α ce scalaire on a :

$$\begin{aligned}(CL)^2 &= (CL)(CL) = C(LC)L \\ &= (LC)CL = \alpha CL\,,\end{aligned}$$

d'où :

$$\begin{aligned}A^2 &= \mathrm{I}_n + (\alpha+2)CL \\ &= \mathrm{I}_n + (\alpha+2)(A - \mathrm{I}_n) \\ &= (\alpha+2)A - (\alpha+1)\mathrm{I}_n\,.\end{aligned}$$

On a donc trouvé un polynôme annulateur de A. Discutons selon α.

– Si $\alpha + 1 \neq 0$, on a :

$$\frac{-1}{\alpha+1} A(A - (\alpha+2)\mathrm{I}_n) = \mathrm{I}_n\,,$$

ce qui prouve que A est inversible, d'inverse :

$$A^{-1} = \frac{-1}{\alpha+1}(A - (\alpha+2)\mathrm{I}_n)\,.$$

– Si $\alpha+1=0$, l'égalité devient

$$A^2=(\alpha+2)A=A\,,$$

donc A n'est pas inversible : sinon, on aurait $A=\mathrm{I}_n$ d'où $CL=0_n$. Mais puisque CL est la matrice dont le terme d'indice (i,j) est $c_i\ell_j$, cela n'est possible que si C et L sont nuls, ce qui est exclu puisque l'on est dans le cas $LC=(-1)$.

Exercice 3.4

• Il est facile de montrer que F est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ (soit en utilisant la caractérisation d'un sous-espace vectoriel, soit en remarquant que c'est le noyau de l' endomorphisme $M\mapsto AMA$).

• D'après **[S3.11]**, on sait qu'il existe $P,\,Q\in\mathrm{GL}_n(\mathbb{K})$ telles que $P^{-1}AQ=J_r$.

En notant alors $B'=Q^{-1}BP=\begin{bmatrix}X&Y\\Z&T\end{bmatrix}$, l'équation $ABA=0$ équivaut à $J_rB'J_r=0$ ce qui donne après un calcul par blocs $X=0$.

Finalement, F est l'ensemble des matrices de la forme $Q\begin{bmatrix}0_r&Y\\Z&T\end{bmatrix}P^{-1}$. Il est donc de dimension n^2-r^2 (car une base en est formée des matrices $QE_{ij}P^{-1}$ avec $(i,j)\notin[\![1\,;r]\!]^2$).

Exercice 3.5

On peut ici raisonner comme dans l'exercice précédent ou l'exemple traité en **[SF3.14]**, en utilisant les matrices J_r, mais nous allons donner une solution différente.

1) Soit $a=\operatorname{rg}A$ et $c=\operatorname{rg}C$. On peut extraire de A une matrice A' inversible d'ordre a et de C une matrice inversible C' d'ordre c. On pourra alors extraire de M une matrice d'ordre $a+c$ de la forme $\begin{bmatrix}A'&0\\B'&C'\end{bmatrix}$; cette matrice étant inversible (utiliser soit le cours sur les matrices triangulaires par blocs, soit le calcul du déterminant par bloc), on en déduit que $\operatorname{rg}M\geqslant a+c$.

2) En posant :

$$A_1=\begin{bmatrix}A&0\\B&0\end{bmatrix}\quad\text{et}\quad A_2=\begin{bmatrix}0&0\\0&C\end{bmatrix},$$

on a $M=A_1+A_2$ donc

$$\operatorname{Im}M=\operatorname{Im}(A_1+A_2)\subset\operatorname{Im}A_1+\operatorname{Im}A_2$$

(en appelant ici image d'une matrice celle de l'application linéaire canoniquement associée).

Puisque, d'après la formule de Grassmann, la dimension de la somme de deux sous-espaces vectoriels est inférieure à la somme des dimensions de ces sous-espaces, on en tire :

$$\operatorname{rg}M\leqslant\operatorname{rg}A_1+\operatorname{rg}A_2\,.$$

$$\text{Or }\dim\operatorname{Im}A_2=\dim\operatorname{Im}\begin{bmatrix}0\\C\end{bmatrix}=\dim\operatorname{Im}C=\operatorname{rg}C\,,$$

et puisque A est inversible, $\operatorname{rg}A_1=\operatorname{rg}A$: en effet on peut extraire de A_1 la matrice inversible A donc $\operatorname{rg}A_1\geqslant\operatorname{rg}A$ et d'autre part $\operatorname{rg}A_1=\operatorname{rg}\begin{bmatrix}A\\B\end{bmatrix}$ est inférieur au nombre de colonnes de A, qui est justement égal à $\operatorname{rg}A$.

On obtient donc $\operatorname{rg}M\leqslant\operatorname{rg}A+\operatorname{rg}C$, d'où l'égalité.

Exercice 3.6

1) • De $u\circ u^2=0$, on tire $\operatorname{Im}u^2\subset\operatorname{Ker}u$. On a aussi, d'après le théorème du rang, $\dim\operatorname{Ker}u=n$.

Considérons alors l'application linéaire $v=u|_{\operatorname{Im}u}$, de $\operatorname{Im}u$ dans E (v est la restriction de u à $\operatorname{Im}u$). On a $\operatorname{Im}v=\operatorname{Im}u^2$ et $\operatorname{Ker}v=\operatorname{Ker}u\cap\operatorname{Im}u$, donc le théorème du rang appliqué à v donne :

$$\operatorname{rg}u=\operatorname{rg}u^2+\dim(\operatorname{Ker}u\cap\operatorname{Im}u)\,.$$

Or $\operatorname{Ker}u\cap\operatorname{Im}u\subset\operatorname{Ker}u$ donc :

$$\dim(\operatorname{Ker}u\cap\operatorname{Im}u)\leqslant n\,,$$

d'où $\operatorname{rg}u^2\geqslant\operatorname{rg}u-n=n$.

Puisque $\operatorname{Im}u^2\subset\operatorname{Ker}u$, on en déduit :

$$\operatorname{Ker}u=\operatorname{Im}u^2\,.$$

• De cette égalité on déduit, d'après le théorème du rang :

$$\dim\operatorname{Im}u=\dim\operatorname{Ker}u^2\,.$$

Puisque $u^2\circ u=0$, on a $\operatorname{Im}u\subset\operatorname{Ker}u^2$, et compte tenu de l'égalité des dimensions :

$$\operatorname{Im}u=\operatorname{Ker}u^2\,.$$

2) Soit S un supplémentaire de $\operatorname{Im} u$; $\dim S = n$; soit $\mathscr{B}_1 = (e_1, \dots, e_n)$ une base de S.
Montrons que la famille :

$$\mathscr{B} = \big(\mathscr{B}_1, u(\mathscr{B}_1), u^2(\mathscr{B}_1)\big)$$

est libre.
Si l'on a des scalaires α_i, β_i et γ_i tels que :

$$\sum_{i=1}^{n} \alpha_i e_i + \sum_{i=1}^{n} \beta_i u(e_i) + \sum_{i=1}^{n} \gamma_i u^2(e_i) = 0\,,$$

alors, en appliquant u^2 à cette égalité, on obtient, puisque $u^3 = 0$:

$$u^2\left(\sum_{i=1}^{n} \alpha_i e_i\right) = 0\,,$$

c'est-à-dire : $\displaystyle\sum_{i=1}^{n} \alpha_i e_i \in \operatorname{Ker} u^2$.

Donc $\displaystyle\sum_{i=1}^{n} \alpha_i e_i \in \operatorname{Im} u$; or $\displaystyle\sum_{i=1}^{n} \alpha_i e_i \in S$, et S et $\operatorname{Im} u$ sont supplémentaires, donc $\displaystyle\sum_{i=1}^{n} \alpha_i e_i = 0$ d'où $\alpha_i = 0$ pour tout $i \in [\![1\,;n]\!]$ puisque la famille $\mathscr{B}_1$ est libre.
Il reste donc :

$$\sum_{i=1}^{n} \beta_i u(e_i) + \sum_{i=1}^{n} \gamma_i u^2(e_i) = 0\,,$$

puis, en appliquant u, on en déduit de la même façon que les β_i sont nuls, et ensuite que les γ_i sont nuls.
Finalement, $\mathscr{B} = \big(\mathscr{B}_1, u(\mathscr{B}_1), u^2(\mathscr{B}_1)\big)$ est une base de E, et il est facile de vérifier que la matrice de u dans cette base est bien de la forme voulue :

$$\begin{array}{ccc} u(\mathscr{B}_1) & u^2(\mathscr{B}_1) & u^3(\mathscr{B}_1) \\ \downarrow & \downarrow & \downarrow \\ \left[\begin{array}{ccc} 0_n & 0_n & 0_n \\ \mathrm{I}_n & 0_n & 0_n \\ 0_n & \mathrm{I}_n & 0_n \end{array}\right] & & \begin{array}{l} \to \mathscr{B}_1 \\ \to u(\mathscr{B}_1) \\ \to u^2(\mathscr{B}_1) \end{array} \end{array}$$

Exercice 3.7

• On peut déjà remarquer que toute matrice antisymétrique est solution ; en effet, si X est antisymétrique, alors les éléments diagonaux de X sont nuls, donc $\operatorname{tr} X = 0$.

• On cherche alors les (éventuelles) solutions qui ne sont pas des matrices antisymétriques.
Si X est une telle matrice, puisque $X + {}^{\mathrm{t}}X = \operatorname{tr}(X) A$, on a forcément $\operatorname{tr} X \neq 0$.
Puisque $X + {}^{\mathrm{t}}X = {}^{\mathrm{t}}(X + {}^{\mathrm{t}}X)$, on en déduit $\operatorname{tr}(X) A = \operatorname{tr}(X)\,{}^{\mathrm{t}}A$, puis nécessairement ${}^{\mathrm{t}}A = A$.

- Si A n'est pas symétrique, il n'y a donc pas de solution de ce type.
- Sinon, en prenant la trace, on a $2\operatorname{tr}(X) = \operatorname{tr}(X)\operatorname{tr}(A)$ d'où nécessairement $\operatorname{tr} A = 2$.
  - Si $\operatorname{tr}(A) \neq 2$, il n'y a pas de solution de ce type.
  - Sinon (on est donc ici dans le cas A symétrique de trace 2) : la partie symétrique de X est nécessairement de la forme λA, donc $X = \lambda A + Y$ avec Y antisymétrique (rappelons que toute matrice est, de façon unique, somme d'une matrice symétrique et d'une matrice antisymétrique). Réciproquement, on vérifie facilement que toute matrice de ce type convient.

• Conclusion.

- Si $\operatorname{tr} A \neq 2$ ou A non symétrique, l'ensemble des solutions est $\mathcal{A}_n(\mathbb{R})$.
- Si A est symétrique de trace 2, l'ensemble des solutions est $\mathcal{A}_n(\mathbb{R}) + \mathbb{R}.A$.

Exercice 3.8

• φ endomorphisme de $\mathcal{M}_n(\mathbb{K})$ est immédiat puisqu'il suffit de vérifier :

$$\varphi(\lambda M + N) = \lambda\varphi(M) + \varphi(N)$$

pour toutes matrices M, $N \in \mathcal{M}_n(\mathbb{K})$ et tout $\lambda \in \mathbb{K}$.

• On écrit $A = \displaystyle\sum_{1 \leqslant k,\ell \leqslant n} a_{k\ell} E_{k\ell}$; pour tout $(i,j) \in [\![1\,;n]\!]^2$ on a alors :

$$\begin{aligned} \varphi(E_{ij}) &= AE_{ij} + E_{ij}A \\ &= \sum_{1 \leqslant k,\ell \leqslant n} a_{k\ell} E_{kl} E_{ij} + \sum_{1 \leqslant k,\ell \leqslant n} a_{k\ell} E_{ij} E_{k\ell} \\ &= \sum_{1 \leqslant k,\ell \leqslant n} a_{k\ell} \delta_{\ell i} E_{kj} + \sum_{1 \leqslant k,\ell \leqslant n} a_{k\ell} \delta_{jk} E_{i\ell} \\ &= \sum_{k=1}^{n} a_{ki} E_{kj} + \sum_{\ell=1}^{n} a_{j\ell} E_{i\ell}\,. \end{aligned}$$

Si l'on considère alors la matrice de φ dans la base canonique $(E_{ij})_{(i,j) \in [\![1;n]\!]^2}$ de

$\mathcal{M}_n(\mathbb{K})$, ses termes diagonaux sont les coordonnées de $\varphi(E_{ij})$ sur E_{ij}, c'est-à-dire, d'après le calcul précédent, les $a_{ii}+a_{jj}$.

Donc :

$$\operatorname{tr}(\varphi)=\sum_{(i,j)\in[\![1;n]\!]^2}(a_{ii}+a_{jj})=2n\operatorname{tr}(A)\,.$$

Exercice 3.9

1) Soit $X\in\mathcal{M}_{n,1}(\mathbb{C})$ un vecteur colonne tel que $AX=0$.
Si $X\neq 0$, on choisit un indice i_0 tel que :

$$\forall\, i\in[\![1\,;n]\!],\ |x_i|\leqslant|x_{i_0}|\ .$$

La ligne d'indice i_0 du système $AX=0$ s'écrit :

$$\sum_{j=1}^{n}a_{i_0j}x_j=0\,,$$

d'où :

$$\sum_{j\neq i_0}a_{i_0j}x_j=-a_{i_0i_0}x_{i_0}\,.$$

En prenant le module et en utilisant l'inégalité triangulaire, on aboutit à :

$$|a_{i_0i_0}|\,|x_{i_0}|\leqslant\sum_{j\neq i_0}|a_{i_0j}|\,|x_j|\ ,$$

et puisque $x_{i_0}\neq 0$ (car $X\neq 0$) :

$$|a_{i_0i_0}|\leqslant\sum_{j\neq i_0}|a_{i_0j}|\underbrace{\frac{|x_j|}{|x_{i_0}|}}_{\leqslant 1}\leqslant\sum_{j\neq i_0}|a_{i_0j}|\ ,$$

ce qui contredit l'hypothèse.

Donc $X=0$ et A est inversible (car $\operatorname{Ker}A=\{0\}$).

2) Pour tout $t\in[0\,;1]$, considérons la matrice A_t obtenue en multipliant par t tous les coefficients de A sauf ceux de la diagonale.
On a $A_0=\operatorname{diag}(a_{11},\dots,a_{nn})$ et $A_1=A$.
De plus, l'application :

$$\varphi:\left\{\begin{array}{ccc}[0\,;1]&\longrightarrow&\mathbb{R}^*\\ t&\longmapsto&\det A_t\end{array}\right.$$

est continue (car polynomiale en t) et à valeurs dans $\mathbb{R}^*$ car toutes les matrices A_t sont encore à diagonale strictement dominante donc inversibles.
Donc φ garde un signe constant, d'après le théorème des valeurs intermédiaires. Puisque $\varphi(0)>0$, on aura $\varphi(1)=\det A>0$.

Exercice 3.10

1) D'après la formule du produit matriciel, le terme d'indice (i,j) de la matrice ${}^tTDT$ est égal à :

$$({}^tTDT)_{ij}=\sum_{k=1}^{n}t_{ki}\varphi(k)t_{kj}\,.$$

Compte tenu de la définition de T il ne reste dans cette somme que :

$$({}^tTDT)_{ij}=\sum_{k\,|\,i\text{ et }k\,|\,j}\varphi(k)\,.$$

Or :

$$k\,|\,i\text{ et }k\,|\,j\iff k\,|\,i\wedge j\,,$$

donc, compte tenu de la formule rappelée par l'énoncé :

$$({}^tTDT)_{ij}=\sum_{k\,|\,i\wedge j}\varphi(k)=i\wedge j\,.$$

2) Ainsi $A={}^tTDT$. Puisque T est triangulaire supérieure avec des coefficients diagonaux égaux à 1, $\det T=\det{}^tT=1$ et ainsi $\det A=\det D$. Finalement :

$$\det A=\prod_{k=1}^{n}\varphi(k)\,.$$

Exercice 3.11

Notons Δ_n le déterminant de A_n.
A_n étant une matrice tridiagonale, on calcule classiquement Δ_n par récurrence ; en développant selon la première colonne, on trouve :

$$\Delta_n=(1+z^2)\Delta_{n-1}-z^2\Delta_{n-2}\,.$$

Cette relation est a priori valable pour $n\geqslant 3$, mais on peut l'étendre à $n=2$ en posant $\Delta_0=1$ (puisque $\Delta_1=1+z^2$ et $\Delta_2=(1+z^2)^2-z^2$).

L'équation caractéristique de cette récurrence linéaire d'ordre 2 est :

$$X^2-(1+z^2)X+z^2=0$$

de racines évidentes 1 et z^2.

Il y a donc deux cas.

- Si $z=\pm1$, il existe α et β tels que $\Delta_n=\alpha n+\beta$. $\Delta_0=1$ et $\Delta_1=2$ donnent $\alpha=\beta=1$ puis $\Delta_n=n+1$.
  Puisque $\Delta_n\neq 0$, la matrice A_n est inversible et son rang est égal à n.

- Sinon, il existe α et β dans $\mathbb{C}$ tels que $\Delta = \alpha z^{2n} + \beta$. $\Delta_0 = 1$ et $\Delta_1 = 1 + z^2$ donnent :
$$\Delta_n = \frac{z^{2n+2} - 1}{z^2 - 1}\,.$$
On en déduit :
  - si $z^{2n+2} \neq 1$, A_n est inversible et son rang est égal à n.
  - si $z^{2n+2} = 1$ (avec $z \neq \pm 1$), A_n n'est pas inversible ; mais on peut extraire de A_n la matrice A_{n-1} qui est inversible puisque $\Delta_{n-1} \neq 0$, donc $\operatorname{rg} A_n = n - 1$.

Exercice 3.12

Soit $V = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ et E_i les vecteurs de la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$.
On a, en utilisant la multilinéarité :
$$\begin{aligned}\Delta &= \det(V + xE_1, V + xE_2, \dots, V + xE_n)\\ &= \underbrace{\det(xE_1, \dots, xE_n)}_{=x^n}\\ &\quad + \det(V, xE_2, \dots, xE_n)\\ &\qquad + \det(xE_1, V, xE_3, \dots, xE_n) + \cdots\\ &\qquad + \det(xE_1, \dots, xE_{n-1}, V)\\ &\quad + \underbrace{\det(V, V, xE_3, \dots, xE_n)}_{=0} + \cdots\end{aligned}$$
et tous les autres déterminants qui apparaissent à la fin sont nuls puisqu'ils comportent au moins deux fois la même colonne V.

Chaque déterminant :
$$\Delta_i = \det(xE_1, \dots, xE_{i-1}, V, xE_{i+1}, \dots, xE_n)$$
pour $i \in [\![1\,;n]\!]$ peut se calculer en le développant selon la i^{e} ligne, mais aussi en utilisant encore la multilinéarité.
En effet, en écrivant $V = \sum_{j=1}^{n} a_j E_j$ on obtient :
$$\begin{aligned}&\det(xE_1, \dots, xE_{i-1}, V, xE_{i+1}, \dots, xE_n)\\ &=\\ &\det(xE_1, \dots, xE_{i-1}, a_i E_i, xE_{i+1}, \dots, xE_n),\end{aligned}$$
les autres déterminants de la somme étant nuls car ils comportent deux colonnes proportionnelles à l'une des colonnes E_j.

Le déterminant obtenu est celui d'une matrice diagonale donc finalement :
$$\Delta_i = a_i x^{n-1},$$
puis : $\Delta = x^n + \left(\sum_{i=1}^{n} a_i\right) x^{n-1}$.

Exercice 4.1

1) On calcule
$$\chi_A = \begin{vmatrix} X-4 & -7 & -6 & 7 \\ -2 & X-3 & -1 & 3 \\ -7 & 1 & X & 7 \\ -4 & -7 & -5 & X+7 \end{vmatrix}.$$
Le calcul semblant un peu compliqué, il faut absolument utiliser des combinaisons linéaires judicieuses sur les rangées.
L'opération $L_1 \leftarrow L_1 - L_4$ donne :
$$\chi_A = \begin{vmatrix} X & 0 & -1 & -X \\ -2 & X-3 & -1 & 3 \\ -7 & 1 & X & 7 \\ -4 & -7 & -5 & X+7 \end{vmatrix}$$
puis par l'opération $C_4 \leftarrow C_4 + C_1$:
$$\chi_A = \begin{vmatrix} X & 0 & -1 & 0 \\ -2 & X-3 & -1 & 1 \\ -7 & 1 & X & 0 \\ -4 & -7 & -5 & X+3 \end{vmatrix}$$
puis par l'opération $L_4 \leftarrow L_4 - (X+3)L_2$:
$$\chi_A = \begin{vmatrix} X & 0 & -1 & 0 \\ -2 & X-3 & -1 & 1 \\ -7 & 1 & X & 0 \\ 2X+2 & -X^2+2 & X-2 & 0 \end{vmatrix}.$$
En développant ensuite selon la dernière colonne :
$$\chi_A = \begin{vmatrix} X & 0 & -1 \\ -7 & 1 & X \\ 2X+2 & 2-X^2 & X-2 \end{vmatrix}.$$
L'opération $C_1 \leftarrow C_1 + XC_3$ donne alors :
$$\chi_A = \begin{vmatrix} 0 & 0 & -1 \\ X^2-7 & 1 & X \\ X^2+2 & 2-X^2 & X-2 \end{vmatrix}.$$
Enfin, en développant selon la première ligne on trouve :
$$\begin{aligned}\chi_A &= -(X^2-7)(2-X^2) + X^2 + 2\\ &= X^4 - 8X^2 + 16 = (X-2)^2(X+2)^2.\end{aligned}$$

2) • On détermine le sous-espace propre associé à la valeur propre 2 en résolvant le système $AV = 2V$, avec $V = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$.

Ce système s'écrit :

$$\begin{cases} 2x + 7y + 6z - 7t = 0 \\ 2x + y + z - 3t = 0 \\ 7x - y - 2z - 7t = 0 \\ 4x + 7y + 5z - 9t = 0 \end{cases}$$

Nous ne détaillons pas la résolution de ce système (faites des combinaisons linéaires judicieuses) ; on trouve que $E_2(f)$ est la droite vectorielle de base (e_1) avec $e_1 = (8,3,2,7)$.

Puisque l'ordre de multiplicité de la valeur propre 2 n'est pas égale à la dimension du sous-espace propre correspondant, on peut déjà conclure que f n'est pas diagonalisable.

• On détermine le sous-espace propre associé à la valeur propre -2 en résolvant le système $AV = -2V$.
Ce système s'écrit :

$$\begin{cases} 6x + 7y + 6z - 7t = 0 \\ 2x + 5y + z - 3t = 0 \\ 7x - y + 2z - 7t = 0 \\ 4x + 7y + 5z - 5t = 0 \end{cases}$$

Après calculs, on trouve que $E_{-2}(f)$ est la droite vectorielle de base (e_3) avec $e_3 = (16,3,-2,15)$

3) Le polynôme caractéristique $\chi_f = (X-2)^2(X+2)^2$ est annulateur de f ; puisque les polynômes $(X-2)^2$ et $(X+2)^2$ sont premiers entre eux, le théorème de décomposition des noyaux donne directement :

$$\mathbb{R}^4 = \operatorname{Ker}(f - 2\mathrm{Id})^2 \oplus \operatorname{Ker}(f + 2\mathrm{Id})^2.$$

4) • On calcule la matrice :

$$A - 2\mathrm{I}_4 = \begin{pmatrix} 2 & 7 & 6 & -7 \\ 2 & 1 & 1 & -3 \\ 7 & -1 & -2 & -7 \\ 4 & 7 & 5 & -9 \end{pmatrix}$$

puis :

$$(A - 2\mathrm{I}_4)^2 = \begin{pmatrix} 32 & -34 & -28 & -14 \\ 1 & -7 & -4 & 3 \\ -30 & 1 & 10 & 31 \\ 21 & -33 & -24 & -3 \end{pmatrix}.$$

Une base de F est formée des vecteurs :

$$e_1 = (8,3,2,7) \text{ et } e_2 = (2,-1,3,1)\,.$$

En effet, puisque :

$$\operatorname{Ker}(f - 2\mathrm{Id}) \subset \operatorname{Ker}(f - 2\mathrm{Id})^2\,,$$

le vecteur e_1, qui est un vecteur propre pour la valeur propre 2, appartient forcément à F ; le second vecteur a été trouvé en écrivant le système $(A - 2\mathrm{I}_4)^2 V = 0$ qui équivaut à

$$\begin{cases} 19x - 6z - 20t = 0 \\ 19y + 10z - 11t = 0 \end{cases}$$

• On calcule la matrice :

$$A + 2\mathrm{I}_4 = \begin{pmatrix} 6 & 7 & 6 & -7 \\ 2 & 5 & 1 & -3 \\ 7 & -1 & 2 & -7 \\ 4 & 7 & 5 & -5 \end{pmatrix}$$

puis :

$$(A + 2\mathrm{I}_4)^2 = \begin{pmatrix} 64 & 22 & 20 & -70 \\ 17 & 17 & 4 & -21 \\ 26 & -7 & 10 & -25 \\ 53 & 23 & 16 & -59 \end{pmatrix}.$$

Le système $(A + 2\mathrm{I}_4)^2 V = 0$ équivaut à

$$\begin{cases} 51x + 18z - 52t = 0 \\ 51y - 6z - 11t = 0 \end{cases}$$

ce qui permet de trouver une base de G formée des vecteurs :

$$e_3 = (16,3,-2,15) \text{ et } e_4 = (2,1,3,3)\,,$$

avec e_3 vecteur propre pour la valeur propre -2.

• Puisque $\mathbb{R}^4 = F \oplus G$, on sait d'après le cours que (e_1,e_2,e_3,e_4) est une base de $\mathbb{R}^4$. Pour écrire la matrice de f dans cette base, il faut exprimer les $f(e_j)$ en fonction des e_i.

– On sait déjà que $f(e_1) = 2e_1$ et $f(e_3) = -2e_3$.

– Puisque F est stable par f, le vecteur $f(e_2)$ appartient à F donc est combinaison linéaire de e_1 et e_2. Les coordonnées dans la base canonique de $f(e_2)$ sont données par le produit matriciel :

$$\begin{pmatrix} 4 & 7 & 6 & -7 \\ 2 & 3 & 1 & -3 \\ 7 & -1 & 0 & -7 \\ 4 & 7 & 5 & -7 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 12 \\ 1 \\ 8 \\ 9 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ -1 \\ 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 8 \\ 3 \\ 2 \\ 7 \end{pmatrix}$$

donc $f(e_2) = e_1 + 2e_2$.

– Les coordonnées dans la base canonique de $f(e_4)$ sont données par le produit matriciel :

$$\begin{pmatrix} 4 & 7 & 6 & -7 \\ 2 & 3 & 1 & -3 \\ 7 & -1 & 0 & -7 \\ 4 & 7 & 5 & -7 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \\ 3 \end{pmatrix}$$

$$= \begin{pmatrix} 12 \\ 1 \\ -8 \\ 9 \end{pmatrix} = -2 \begin{pmatrix} 2 \\ 1 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 16 \\ 3 \\ -2 \\ 15 \end{pmatrix}$$

donc $f(e_4) = e_3 - 2e_4$.

Il en résulte que la matrice de f dans la nouvelle base est :

$$A' = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

✓ On a donc trigonalisé la matrice A.

Exercice 4.2

On remarque d'emblée que $A = \mathrm{I}_n + aJ$ où J est la matrice :

$$J = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}$$

et on est donc ramené à la diagonalisation de J.

On constate alors que J est symétrique réelle, donc diagonalisable. De plus, $\operatorname{rg} J = 3$, puisque les colonnes C_2 à C_{n-1} sont égales, et que les trois colonnes (C_1, C_2, C_n) sont visiblement indépendantes.

On peut déjà en déduire que $\operatorname{Ker} J$ est de dimension $n-3$, donc 0 sera valeur propre d'ordre $n-3$ (0 n'est pas valeur propre si $n = 3$). Mais il reste trois valeurs propres à déterminer...

On pourrait bien sûr, si on note α, β et γ ces trois valeurs propres, résoudre le système :

$$\begin{cases} \alpha + \beta + \gamma &= \operatorname{tr} J \\ \alpha^2 + \beta^2 + \gamma^2 &= \operatorname{tr} J^2 \\ \alpha^3 + \beta^3 + \gamma^3 &= \operatorname{tr} J^3 \end{cases}$$

(en effet, J étant semblable à $\operatorname{diag}(\alpha,\beta,\gamma,0,\ldots)$, J^k est semblable à $\operatorname{diag}(\alpha^k,\beta^k,\gamma^k,0,\ldots)$), mais le problème de cette méthode, outre les calculs, est qu'elle donnera seulement les valeurs propres, et il faudra faire de nouveaux calculs pour obtenir les sous-espaces propres.

Nous préfèrerons donc la méthode qui permet d'obtenir les deux résultats en même temps : on cherche les réels λ pour lesquels il existe $V \in \mathcal{M}_{n,1}(\mathbb{R})$, non nul, tel que $JV = \lambda V$.

Cela conduit au système :

$$(S) \begin{cases} x_2 + \cdots + x_n &= \lambda x_1 \\ x_1 + x_n &= \lambda x_2 \\ \vdots &= \vdots \\ x_1 + x_n &= \lambda x_{n-1} \\ x_1 + \cdots x_{n-1} &= \lambda x_n \end{cases}$$

– Le cas $\lambda = 0$ conduit au système :

$$\begin{cases} x_1 + \cdots + x_{n-1} = 0 \\ x_1 = 0 \\ x_n = 0 \end{cases}$$

qui est bien un système d'équations d'un sous-espace vectoriel de dimension $n-3$ (intersection de 3 hyperplans linéairement indépendants).

– Pour $\lambda \neq 0$, les équations centrales de (S) donnent $x_2 = \ldots = x_{n-1}$ et le reste du système équivaut à :

$$\begin{cases} (n-2)x_2 + x_n = \lambda x_1 \\ x_1 + x_n = \lambda x_2 \\ x_1 + (n-2)x_2 = \lambda x_n \end{cases}$$

soit à :

$$\begin{cases} -\lambda x_1 + (n-2)x_2 + x_n = 0 \\ x_1 - \lambda x_2 + x_n = 0 \\ x_1 + (n-2)x_2 - \lambda x_n = 0 \end{cases}$$

En soustrayant la première équation à la dernière, on obtient $(\lambda+1)(x_1 - x_n) = 0$, ce qui conduit à distinguer deux cas.

– Pour $\lambda = -1$, le système précédent devient :

$$\begin{cases} x_1 + (n-2)x_2 + x_n = 0 \\ x_1 + x_2 + x_n = 0 \end{cases}$$

d'où $x_2 = 0$ et $x_1 = -x_n$.

Ainsi, -1 est valeur propre, et le sous-espace propre associé est la droite vectorielle de base ${}^{\mathrm{t}}\begin{pmatrix} 1 & 0 & \ldots & 0 & -1 \end{pmatrix}$.

– Si $\lambda \neq -1$ alors le système précédent équivaut à :

$$\begin{cases} (1-\lambda)x_1 + (n-2)x_2 = 0 \\ 2x_1 - \lambda x_2 = 0 \\ x_1 - x_n = 0 \end{cases}$$

ce qui implique en particulier

$$\left((1-\lambda)\frac{\lambda}{2} + n - 2\right) x_2 = 0\,.$$

Le cas $x_2 = 0$ étant exclu car il conduirait à $V = 0$, on a donc forcément $(1-\lambda)\dfrac{\lambda}{2} + n - 2 = 0$. Cette équation du second degré en λ a deux racines réelles λ_1 et λ_2 (qu'il est inutile d'exprimer pour l'instant) ; chacune de ces racines est une valeur propre de J, et le sous-espace propre associé est la droite de base ${}^{\mathrm{t}}\left(\frac{\lambda_i}{2} \quad 1 \quad \dots \quad 1 \quad \frac{\lambda_i}{2}\right)$ (obtenu pour $x_2 = 1$).

Conclusion.

– Pour $n = 3$, la matrice J possède deux valeurs propres : -1, double (car ici -1 est aussi racine de l'équation $(1-\lambda)\dfrac{\lambda}{2} + n - 2 = 0$), de sous-espace propre associé le plan d'équation $x + y + z = 0$, et 2, simple, de sous-espace propre associé la droite de base ${}^{\mathrm{t}}\begin{pmatrix}1 & 1 & 1\end{pmatrix}$.

– Pour $n > 3$, J possède 4 valeurs propres :

  – 0, d'ordre $n-3$; le sous-espace propre associé est le sous-espace vectoriel dont un système d'équations est :

$$\begin{cases} x_1 + \cdots + x_{n-1} = 0 \\ x_1 = x_n = 0 \end{cases}$$

  – -1, simple ; le sous-espace propre associé est la droite de base le vecteur ${}^{\mathrm{t}}\begin{pmatrix}1 & 0 & \dots & 0 & -1\end{pmatrix}$.

  – $\lambda_i = \dfrac{1 \pm \sqrt{8n-15}}{2}$, simples ;
  chaque sous-espace propre associé est la droite de base le vecteur ${}^{\mathrm{t}}\left(\frac{\lambda_i}{2} \quad 1 \quad \dots \quad 1 \quad \frac{\lambda_i}{2}\right)$.

Enfin, pour terminer la résolution de l'exercice, les valeurs propres de $A = \mathrm{I}_n + aJ$ sont égales à $a\lambda + 1$ où $\lambda \in \operatorname{Sp} J$, et les sous-espaces propres de A sont les mêmes que ceux de J dans le cas $a \neq 0$.

Exercice 4.3

On vérifie facilement (cela est implicite dans l'énoncé), que u est bien un endomorphisme de $\mathbb{R}[X]$.
On passe donc au point essentiel, celui de la détermination des éléments propres.
Il faut donc chercher les réels λ pour lesquels il existe $P \neq 0$ tel que $u(P) = \lambda P$, ce qui équivaut à :

$$(aX + b - \lambda)P + X(1-X)P' = 0\,.$$

On peut résoudre cette équation par le calcul (en posant $P = \sum_{k=0}^{n} p_k X^k$), mais cela conduit à des systèmes d'équations assez délicats à résoudre. Il est bien sûr préférable d'utiliser la théorie des équations différentielles.
En considérant la fonction polynôme associée à P, on est amené à rechercher les solutions polynomiales de l'équation différentielle :

$$x(1-x)y' + (ax + b - \lambda)y = 0 \quad (E)$$

Pour résoudre cette équation, on se place sur un intervalle où le coefficient de y' ne s'annule pas, donc sur $]0\,;1[$ par exemple. Dans le cas présent, cette restriction n'est pas gênante puisque l'on cherche les solutions polynomiales, et qu'un polynôme est entièrement connu si l'on connait sa valeur sur un ensemble infini.

L'équation (E) équivaut alors à :

$$y' + \frac{ax + b - \lambda}{x(1-x)}y = 0\,.$$

D'après le cours (voir **[S17.6]**), l'ensemble des solutions de cette équation est l'ensemble des fonctions de la forme :

$$x \mapsto C \exp\left(-\int \frac{ax + b - \lambda}{x(1-x)}\,\mathrm{d}x\right).$$

On effectue une décomposition en éléments simples :

$$\frac{ax + b - \lambda}{x(1-x)} = \frac{b - \lambda}{x} + \frac{a + b - \lambda}{1 - x}\,.$$

Donc, compte tenu du fait que $x \in\ ]0\,;1[$:

$$\int \frac{ax + b - \lambda}{x(1-x)}\,\mathrm{d}x = (b - \lambda)\ln x - (a + b - \lambda)\ln(1-x)\,,$$

d'où l'on tire les solutions de (E) :

$$\begin{aligned} y &= C\exp\big(-(b-\lambda)\ln x \\ &\qquad +(a+b-\lambda)\ln(1-x)\big) \\ &= Cx^{\lambda-b}(1-x)^{a+b-\lambda}\,. \end{aligned}$$

Une telle solution est une solution polynomiale si et seulement si :

$$\lambda - b \in \mathbb{N} \quad \text{et} \quad a+b-\lambda \in \mathbb{N}$$

ce qui équivaut à dire que λ est un entier compris (au sens large) entre b et $a+b$.

Les valeurs propres de u sont donc les entiers $b+i$ avec $i \in [\![0\,;a]\!]$.
De plus, par cette méthode, on a aussi obtenu les vecteurs propres : le sous-espace propre associé à la valeur propre λ est la droite vectorielle $\mathrm{Vect}\left(X^{\lambda-b}(1-X)^{a+b-\lambda}\right)$, donc celui associé à la valeur propre $b+i$ avec $i \in [\![0\,;a]\!]$ est $\mathrm{Vect}\left(X^i(1-X)^{a-i}\right)$.

Exercice 4.4

Posons $F = u(f)$.
On a, pour tout $x \in [0\,;1]$:

$$F(x) = \int_0^x tf(t)\,\mathrm{d}t + x\int_x^1 f(t)\,\mathrm{d}t\,.$$

Cela montre que F est de classe $\mathscr{C}^1$, (donc, au passage, u est bien un endomorphisme de E), et permet de calculer sa dérivée :

$$\forall\, x \in [0\,;1],\ F'(x) = \int_x^1 f(t)\,\mathrm{d}t\,,$$

donc F' est de classe $\mathscr{C}^1$, puis $F''(x) = -f(x)$.

Pour déterminer les éléments propres de u, on cherche les $f \in E$, non nulles telles que $F = u(f) = \lambda f$.

• Si $\lambda = 0$, on a $F = 0$ d'où $f = -F'' = 0$: donc 0 n'est pas valeur propre.

• Sinon f est de classe $\mathscr{C}^2$ et vérifie : $\lambda f'' + f = 0$. La résolution de cette équation différentielle linéaire homogène du second ordre à coefficients constants est immédiate. Puisque $f(0) = 0$, on obtient, nécessairement :

$$f(x) = \begin{cases} C\sin\dfrac{x}{\sqrt{\lambda}} & \text{si } \lambda > 0 \\[2ex] C\,\mathrm{sh}\dfrac{x}{\sqrt{-\lambda}} & \text{si } \lambda < 0 \end{cases}$$

où C est une constante réelle.
De plus, l'expression de F' montre que $F'(1) = 0$, donc, puisque $F' = \lambda f'$, on a aussi $f'(1) = 0$.

– Si $\lambda > 0$, une solution est de la forme $f\colon x \mapsto C\sin\frac{x}{\sqrt{\lambda}}$. On a alors $f'(1) = \frac{C}{\sqrt{\lambda}}\cos\frac{1}{\sqrt{\lambda}}$.
Puisque $C \neq 0$ (sinon $f = 0$ ce qui est exclu), on doit donc avoir $\frac{1}{\sqrt{\lambda}} \in \pi\mathbb{Z} + \frac{\pi}{2}$, soit

$$\lambda = \frac{1}{\left(k+\frac{1}{2}\right)^2\pi^2} \text{ avec } k \in \mathbb{Z}\,,$$

mais il suffit de prendre $k \in \mathbb{N}$ pour avoir des valeurs distinctes.

– Si $\lambda < 0$, alors $f(x) = C\,\mathrm{sh}\frac{x}{\sqrt{-\lambda}}$, et on a $f'(x) = \frac{C}{\sqrt{-\lambda}}\,\mathrm{ch}\frac{x}{\sqrt{-\lambda}}$, donc on ne pourra pas avoir $f'(1) = 0$: le problème n'a pas de solution dans ce cas.

Puisque nous avons procédé par conditions nécessaires (analyse), il faut vérifier que, réciproquement (synthèse), les fonctions de la forme $x \mapsto C\sin\left(\left(k+\frac{1}{2}\right)\pi x\right)$ conviennent, ce que nous vous laissons le soin de faire.

• En conclusion, le spectre de u est

$$\mathrm{Sp}\,u = \left\{\frac{1}{\left(k+\frac{1}{2}\right)^2\pi^2}\,,\ k \in \mathbb{N}\right\}\,,$$

le sous-espace propre associé à la valeur propre $\frac{1}{\left(k+\frac{1}{2}\right)^2\pi^2}$ $(k \in \mathbb{N})$ étant

$$E_{\lambda_k} = \left\{x \mapsto C\sin\left(\left(k+\frac{1}{2}\right)\pi x\right),\ C \in \mathbb{R}\right\}.$$

Exercice 4.5

• $(i) \implies (ii)$ est immédiat, car si (i) est vérifiée, $\mathrm{Im}(u-\lambda\mathrm{Id}_E)$ est un supplémentaire de $E_\lambda(u)$ stable par u.

• $(ii) \implies (iii)$. Soit F un sous-espace vectoriel stable par u et supplémentaire de $E_\lambda(u)$.
En notant $r = \dim E_\lambda(u)$, la matrice de u dans une base adaptée à la somme directe $E = E_\lambda(u) \oplus F$ est de la forme :

$$\begin{bmatrix} \lambda\mathrm{I}_r & 0 \\ 0 & A \end{bmatrix}\,,$$

où A est la matrice de l'endomorphisme u_F induit par u sur F. Donc $\chi_u(X) = (X-\lambda)^r \chi_{u_F}(X)$, et puisque λ n'est pas valeur propre de u_F, l'ordre de multiplicité de λ est égal à $r = \dim E_\lambda(u)$.

• $(iii) \implies (iv)$. Notons toujours $r = \dim E_\lambda(u)$. L'hypothèse implique :

$$\chi_u(X) = (X-\lambda)^r Q(X)\,,$$

où $Q(\lambda) \neq 0$, et le théorème de décomposition des noyaux fournit la décomposition :

$$E = \operatorname{Ker}(u-\lambda \mathrm{Id}_E)^r \oplus \operatorname{Ker} Q(u)\,.$$

Si l'on note u' l'endomorphisme induit par u sur $\operatorname{Ker}(u-\lambda\mathrm{Id}_E)^r$ et u'' celui sur $\operatorname{Ker} Q(u)$, on a $\chi_u = \chi_{u'} \cdot \chi_{u''}$. Or u'' ne peut admettre λ pour valeur propre puisque $Q(u'') = 0$ et $Q(\lambda) \neq 0$; donc $\chi_{u'} = (X-\lambda)^r$, ce qui implique que :

$$\dim \operatorname{Ker}(u-\lambda\mathrm{Id}_E)^r = \deg \chi_{u'} = r$$

et puisque $r = \dim \operatorname{Ker}(u-\lambda\mathrm{Id}_E)$, on a $\operatorname{Ker}(u-\lambda\mathrm{Id}_E)^r = \operatorname{Ker}(u-\lambda\mathrm{Id}_E)$.
Finalement, $E = E_\lambda(u) \oplus \operatorname{Ker} Q(u)$, et la matrice de u dans une base associée à cette décomposition est de la forme :

$$\begin{bmatrix} \lambda \mathrm{I}_r & 0 \\ 0 & A'' \end{bmatrix},$$

où A'' est la matrice de u''. Le polynôme minimal $\pi_{u''}$ de u'' ne s'annule pas en λ puisque $\pi_{u''}$ divise Q. L'écriture matricielle ci-dessus montre alors que le polynôme minimal de u est égal à $(X-\lambda)\pi_{u''}$; λ en est bien une racine simple.

• $(iv) \implies (i)$. Si (iv) est vérifiée, le polynôme minimal de u est de la forme $(X-\lambda)Q$ où $Q(\lambda) \neq 0$. Le théorème de décomposition des noyaux fournit alors :

$$E = E_\lambda(u) \oplus \operatorname{Ker} Q(u)\,,$$

et puisque $\operatorname{Ker} Q(u)$ est stable par u, on a le résultat voulu.

• Montrons pour finir que, dans ces conditions, $\operatorname{Im}(u-\lambda\mathrm{Id}_E)$ est le seul supplémentaire de $E_\lambda(u)$ stable par u.

En effet, si F est tel que $E = E_\lambda(u) \oplus F$, on a, puisque $(u-\lambda\mathrm{Id}_E)(E_\lambda(u)) = \{0\}$,

$$\operatorname{Im}(u-\lambda\mathrm{Id}_E) = (u-\lambda\mathrm{Id}_E)(F) \subset F\,.$$

L'égalité des dimensions à l'aide du théorème du rang permet alors de conclure $F = \operatorname{Im}(u-\lambda\mathrm{Id}_E)$.

Exercice 4.6

En considérant le vecteur colonne $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ u_{n+2} \end{pmatrix}$, la relation de récurrence de l'énoncé peut s'écrire sous la forme :

$$X_{n+1} = AX_n\,,$$

avec :

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 4\cos\theta - 1 & 2(\cos\theta - 1) \end{pmatrix}.$$

Le polynôme caractéristique χ_A de cette matrice se calcule aisément ; il vaut

$$X^3 - 2(\cos\theta - 1)X^2 + (1 - 4\cos\theta)X + 2\,.$$

On remarque que -2 est racine « évidente » de ce polynôme, de sorte que l'on peut factoriser :

$$\chi_A = (X+2)(X^2 - 2X\cos\theta + 1)\,.$$

χ_A possède ainsi trois racines distinctes dans $\mathbb{C}$: -2, $\mathrm{e}^{\mathrm{i}\theta}$ et $\mathrm{e}^{-\mathrm{i}\theta}$.
A est donc diagonalisable et il existe $P \in \mathrm{GL}_3(\mathbb{C})$ telle que :

$$A = PDP^{-1} \text{ avec } D = \operatorname{diag}(-2, \mathrm{e}^{\mathrm{i}\theta}, \mathrm{e}^{-\mathrm{i}\theta}).$$

On aura alors $A^n = PD^nP^{-1}$ et puisque $X_n = A^nX_0$ on peut dire sans avoir besoin de calculer P ni P^{-1} (ouf !) qu'il existe des complexes a, b, c tels que :

$$\forall\, n \in \mathbb{N},\ u_n = a(-2)^n + b\mathrm{e}^{\mathrm{i}n\theta} + c\mathrm{e}^{-\mathrm{i}n\theta}\,,$$

ou encore, puisque ici la suite est à valeurs réelles, qu'il existe des réels α, β, γ tels que :

$$\forall\, n \in \mathbb{N},\ u_n = \alpha(-2)^n + \beta\cos n\theta + \gamma \sin n\theta\,.$$

Les conditions initiales de l'énoncé amènent au système :

$$\begin{cases} \alpha + \beta & = 1 + \cos\theta \\ -2\alpha + \beta\cos\theta + \gamma\sin\theta & = -1 \\ 4\alpha + \beta\cos 2\theta + \gamma\sin 2\theta & = 4 + \cos\theta \end{cases}$$

qui a comme solutions :

$$\alpha = 1\ ,\ \beta = \cos\theta\ ,\ \gamma = \sin\theta\,.$$

En conclusion on a :

$$\forall\, n \in \mathbb{N},\ u_n = (-2)^n + \cos(n-1)\theta\,.$$

Exercice 4.7

Un polynôme annulateur de A est $P = X^3 - 3X + 4$. Pour obtenir des renseignements sur ses racines (donc sur les valeurs propres possibles de A), on étudie rapidement la fonction $f : x \mapsto x^3 - 3x + 4$. On obtient le tableau suivant :

x	$-\infty$		-1		1		$+\infty$
$f'(x)$		$+$	0	$-$	0	$+$	
$f(x)$	$-\infty$	↗	6	↘	2	↗	$+\infty$

Ce tableau, ainsi que le théorème de la bijection, montrent que f possède une et une seule racine réelle λ, avec $\lambda < -1$. Le polynôme P étant à coefficients réels possède donc aussi deux racines complexes conjuguées μ et $\overline{\mu}$, et $\mathrm{Sp}(u) \subset \{\lambda, \mu, \overline{\mu}\}$.

Si on note p l'ordre de multiplicité de la valeur propre λ dans le polynôme caractéristique de A, et q celui de la valeur propre μ (c'est aussi l'ordre de $\overline{\mu}$ puisque χ_A est à coefficients réels), avec éventuellement p ou q nul, on aura :

$$\det A = \lambda^p \mu^q \overline{\mu}^q = \lambda^p (|\mu|^2)^q \,,$$

donc $\det A$ est du signe de $(-1)^p$, et puisque $n = p + 2q$, $\det A$ est du signe de $(-1)^n$.

Exercice 4.8

Remarque : dans tout l'exercice, on confondra les matrices avec les endomorphismes qui leur sont canoniquement associés.

A et B étant symétriques réelles sont diagonalisables dans $\mathcal{M}_n(\mathbb{R})$.
Notons $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de A. Puisque A est semblable à $\mathrm{diag}(\lambda_1, \ldots, \lambda_1, \ldots, \lambda_p, \ldots, \lambda_p)$ (chaque valeur propre étant répétée autant de fois que son ordre de multiplicité), A^3 est semblable à $\mathrm{diag}(\lambda_1^3, \ldots, \lambda_1^3, \ldots, \lambda_p^3, \ldots, \lambda_p^3)$.

L'application $x \mapsto x^3$ étant injective sur $\mathbb{R}$, cela prouve que les valeurs propres de A^3 sont exactement les λ_i^3 (avec le même ordre de multiplicité).
De même, si l'on note $\mu_1, \ldots, \mu_q$ les valeurs propres distinctes de B, celles de B^3 sont exactement les μ_i^3. Puisque $A^3 = B^3$, les listes $(\lambda_1^3, \ldots, \lambda_p^3)$ et $(\mu_1^3, \ldots, \mu_q^3)$ sont les mêmes, ce qui implique, toujours par injectivité de l'application $x \mapsto x^3$, que $p = q$ et que les valeurs propres de B sont les mêmes que celles de A.

Pour tout $i \in [\![1\,;p]\!]$, on a :

$$\mathrm{Ker}(B - \lambda_i \mathrm{I}_n) \subset \mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n) \,,$$

car :

$$\begin{aligned} V \in \mathrm{Ker}(B - \lambda_i \mathrm{I}_n) &\implies BV = \lambda_i V \\ &\implies B^3 V = \lambda_i^3 V \\ &\implies A^3 V = \lambda_i^3 V \,. \end{aligned}$$

D'autre part, puisque A^3 et B commutent, pour tout $i \in [\![1\,;p]\!]$, $\mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n)$ est stable par B. Il résulte alors d'un théorème du cours que la restriction de B à $\mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n)$ est diagonalisable et puisque sa seule valeur propre ne peut être que λ_i (car si $V \in \mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n)$, $BV = \lambda V \implies A^3 V = \lambda^3 V = \lambda_i^3 V$ d'où $\lambda_i^3 = \lambda^3$ puis $\lambda = \lambda_i$), cette restriction est l'homothétie de rapport λ_i, c'est-à-dire que $\mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n) \subset \mathrm{Ker}(B - \lambda_i \mathrm{I}_n)$.
Ainsi, tout $i \in [\![1\,;p]\!]$:

$$\mathrm{Ker}(B - \lambda_i \mathrm{I}_n) = \mathrm{Ker}(A^3 - \lambda_i^3 \mathrm{I}_n) \,.$$

On démontre de même :

$$\mathrm{Ker}(A - \lambda_i \mathrm{I}_n) = \mathrm{Ker}(B^3 - \lambda_i^3 \mathrm{I}_n) \,,$$

donc pour tout $i \in [\![1\,;p]\!]$:

$$\mathrm{Ker}(A - \lambda_i \mathrm{I}_n) = \mathrm{Ker}(B - \lambda_i \mathrm{I}_n) \,,$$

et ces sous-espaces sont supplémentaires.
A et B sont donc diagonalisables dans la même base et ont les mêmes valeurs propres : elles sont égales.

Exercice 4.9

1) Le polynôme caractéristique de B est donné par :

$$\forall \lambda \in \mathbb{C} \,,\ \chi_B(\lambda) = \begin{vmatrix} \lambda \mathrm{I}_n & -\mathrm{I}_n \\ -A & \lambda \mathrm{I}_n \end{vmatrix} \,.$$

En effectuant, pour $\lambda \neq 0$, les opérations élémentaires sur les colonnes :
$C_{n+i} \leftarrow C_{n+i} + \dfrac{1}{\lambda} C_i$ pour i variant de 1 à n, on obtient pour tout $\lambda \in \mathbb{C}^*$,

$$\chi_B(\lambda) = \begin{vmatrix} \lambda \mathrm{I}_n & 0 \\ -A & \lambda \mathrm{I}_n - \dfrac{1}{\lambda} A \end{vmatrix}$$

ce qui est le déterminant d'une matrice triangulaire inférieure par blocs, donc, pour tout λ non nul :

$$\begin{aligned}\chi_B(\lambda) &= \det(\lambda \mathrm{I}_n)\det\left(\lambda \mathrm{I}_n - \frac{1}{\lambda}A\right)\\ &= \lambda^n \det\left(\frac{1}{\lambda}\left(\lambda^2 \mathrm{I}_n - A\right)\right)\\ &= \lambda^n \left(\frac{1}{\lambda}\right)^n \det(\lambda^2 \mathrm{I}_n - A) = \chi_A(\lambda^2).\end{aligned}$$

Cette égalité entre deux polynômes étant valable pour une infinité de valeurs de λ est en fait valable pour tout λ donc :

$$\forall\, \lambda \in \mathbb{C}\ ,\ \chi_B(\lambda) = \chi_A(\lambda^2)\,.$$

2) Une solution pour résoudre cette question est de déterminer les éléments propres de la matrice, en utilisant des calculs par blocs, comme dans l'exemple de **[SF4.10]** page 303.

Nous allons faire différemment, en espérant vous convaincre que l'utilisation des polynômes de matrices, lorsqu'elle est bien maitrisée (et lorsqu'elle est possible !), est une méthode très élégante.

On a :

$$B = \begin{bmatrix} 0 & \mathrm{I}_n \\ A & 0 \end{bmatrix} \quad \text{et} \quad B^2 = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$$

donc, par récurrence sur k, on obtient :

$$\forall\, k \in \mathbb{N}\ ,\ B^{2k} = \begin{bmatrix} A^k & 0 \\ 0 & A^k \end{bmatrix}\ ,$$

puis, si P est un polynôme quelconque, on aura par linéarité :

$$P(B^2) = \begin{bmatrix} P(A) & 0 \\ 0 & P(A) \end{bmatrix} \quad (1)$$

– Supposons dans un premier temps A diagonalisable et inversible. Il existe un polynôme P scindé à racines simples tel que $P(A) = 0$.

  Si 0 est racine (simple, par hypothèse) de P, on peut écrire $P = XQ$ avec Q scindé à racines simples et n'admettant pas 0 pour racine. L'égalité $P(A) = 0$ implique alors $AQ(A) = 0$ d'où $Q(A) = 0$ puisque A est inversible.

  On peut donc supposer, quitte à remplacer P par Q, que P n'admet pas 0 pour racine. Quitte à multiplier P par une constante convenable, on peut aussi supposer P unitaire. On peut donc écrire :

$$P = \prod_{i=1}^{p}(X - \lambda_i)$$

avec les $\lambda_i \in \mathbb{C}^*$ distincts deux à deux. D'après (1), la relation $P(A) = 0$ implique $P(B^2) = 0$ donc :

$$\prod_{i=1}^{p}(B^2 - \lambda_i I_{2n}) = 0_{2n}\,.$$

Chaque λ_i étant un complexe non nul possède exactement deux racines complexes distinctes et non nulles, que nous noterons $\pm\mu_i$: $\mu_i^2 = \lambda_i$. Alors, pour tout $i \in [\![1\,;p]\!]$,

$$\begin{aligned}B^2 - \lambda_i I_{2n} &= B^2 - \mu_i^2 I_{2n}\\ &= (B - \mu_i I_{2n})(B + \mu_i I_{2n})\end{aligned}$$

et on en déduit :

$$\prod_{i=1}^{p}(B - \mu_i I_{2n})(B + \mu_i I_{2n}) = 0_{2n}\,.$$

Cela signifie que le polynôme $\prod_{i=1}^{p}(X - \mu_i)(X + \mu_i)$ est annulateur de B ; ce polynôme étant scindé à racines simples puisque les $\pm\mu_i$ sont tous distincts, on a donc démontré que :

$$\begin{aligned}&A \text{ diagonalisable et inversible}\\ &\Longrightarrow B \text{ diagonalisable.}\end{aligned}$$

– Supposons maintenant B diagonalisable.

  – B^2 est alors aussi diagonalisable puisque, si B est semblable à $\mathrm{diag}(\lambda_1, \ldots, \lambda_{2n})$, B^2 est semblable à $\mathrm{diag}(\lambda_1^2, \ldots, \lambda_{2n}^2)$.

    Il existe donc un polynôme P scindé à racines simples tel que $P(B^2) = 0$, et la relation (1) implique alors $P(A) = 0$.

    Il en résulte que A est aussi diagonalisable.

  – B étant diagonalisable, il existe un polynôme Q scindé à racines simples tel que $Q(B) = 0$.

    Les calculs précédents montrent qu'il faut séparer les puissances paires et impaires de B.

    Pour cela, on écrit le polynôme Q comme somme d'un polynôme pair (celui formé des termes en X^{2k}) et

d'un polynôme impair (celui formé des termes en X^{2k+1}), soit :

$$Q(X) = R(X^2) + XS(X^2)$$

avec R, S polynômes.

On aura alors, compte tenu de (1) :

$$\begin{aligned} Q(B) &= R(B^2) + BS(B^2) \\ &= \begin{bmatrix} R(A) & 0 \\ 0 & R(A) \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & \mathrm{I}_n \\ A & 0 \end{bmatrix} \times \begin{bmatrix} S(A) & 0 \\ 0 & S(A) \end{bmatrix} \\ &= \begin{bmatrix} R(A) & S(A) \\ AS(A) & R(A) \end{bmatrix}. \end{aligned}$$

L'égalité $Q(B) = 0$ implique donc $R(A) = S(A) = 0$, c'est-à-dire que R et S sont annulateurs de A.

Si, par l'absurde, A n'était pas inversible, 0 serait valeur propre de A donc racine de R et S (car les valeurs propres de A sont incluses dans l'ensemble des racines d'un polynôme annulateur). 0 serait alors racine au moins double de $R(X^2)$ et de $S(X^2)$, donc de $Q(X)$, ce qui est impossible puisque l'on a supposé Q à racines simples.

Ainsi, 0 n'est pas valeur propre de A ; il en résulte que A est inversible.

Finalement on a bien montré que :

B diagonalisable $\Longrightarrow$ A diagonalisable et inversible.

3) Pour répondre à cette question, il suffit de reprendre la démonstration précédente, en utilisant simplement le fait que seuls les nombres réels positifs possèdent une racine carrée réelle.

On obtient le résultat suivant :

lorsque $A \in \mathcal{M}_n(\mathbb{R})$, la matrice B est diagonalisable dans $\mathcal{M}_{2n}(\mathbb{R})$ si et seulement si A est diagonalisable dans $\mathcal{M}_n(\mathbb{R})$ et a toutes ses valeurs propres strictement positives.

Exercice 4.10

Déjà, il est facile de vérifier que f_A et g_A sont bien des endomorphismes de $\mathcal{M}_n(\mathbb{K})$.

Par exemple :

$\forall \lambda \in \mathbb{K}$, $\forall (M,N) \in \mathcal{M}_n(\mathbb{K})^2$,

$$\begin{aligned} f_A(\lambda M + N) &= A(\lambda M + N) \\ &= \lambda f_A(M) + f_A(N). \end{aligned}$$

1) • Supposons A inversible.

Il est alors facile de démontrer que $f_{A^{-1}} \circ f_A = \mathrm{Id}_{\mathcal{M}_n(\mathbb{K})}$, puisque, pour toute $M \in \mathcal{M}_n(\mathbb{K})$,

$$f_{A^{-1}}[f_A(M)] = A^{-1}AM = M.$$

f_A possède donc un inverse à gauche qui est $f_{A^{-1}}$, et comme il s'agit d'un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie, c'est aussi son inverse à droite.

• Supposons f_A inversible.

f_A est en particulier surjective, donc la matrice I_n a un antécédent par f_A : $\exists\, M \in \mathcal{M}_n(\mathbb{K})$ tel que $AM = \mathrm{I}_n$. Cela signifie que A est inversible à droite, donc inversible, d'inverse M.

2) a) • Soit $\lambda \in \mathrm{Sp}(A)$. Cela signifie qu'il existe une matrice colonne $X \in \mathcal{M}_{n,1}(\mathbb{K})$, non nulle, telle que $AX = \lambda X$.

Si l'on considère alors la matrice écrite par blocs $M = \begin{bmatrix} X & X & \dots & X \end{bmatrix}$ (dont toutes les colonnes sont égales à X), on a $AM = \begin{bmatrix} AX & \dots & AX \end{bmatrix} = \lambda M$, soit $f_A(M) = \lambda M$. M étant non nulle, cela signifie que λ est aussi valeur propre de f_A.

• Soit $\lambda \in \mathrm{Sp}(f_A)$. Cela signifie qu'il existe une matrice $M \in \mathcal{M}_n(\mathbb{K})$, non nulle, telle que $AM = \lambda M$. En écrivant là encore la matrice M par blocs sous la forme $M = \begin{bmatrix} C_1 & C_2 & \dots & C_n \end{bmatrix}$ (les C_j étant les colonnes de la matrice), on a

$$\begin{aligned} f_A(M) &= \begin{bmatrix} AC_1 & AC_2 & \dots & AC_n \end{bmatrix} \\ &= \lambda M \end{aligned}$$

d'où $AC_j = \lambda C_j$ pour tout $j \in [\![1\,;n]\!]$. La matrice M étant non nulle, il existe au moins une colonne C_j non nulle, ce qui prouve que λ est bien valeur propre de A.

En conclusion :

f_A et A ont mêmes valeurs propres.

b) Nous proposons deux méthodes, parmi les plus importantes, pour étudier la diagonalisabilité de f_A.

• *Utilisation d'un polynôme annulateur*

Il est facile de remarquer ici que :

$$\forall\, M \in \mathcal{M}_n(\mathbb{K}),$$

$$f_A \circ f_A(M) = A^2M = f_{A^2}(M)$$

et plus généralement, que, pour tout $k \in \mathbb{N}$: $f_{A^k} = (f_A)^k$ (cette formule est immédiate pour $k = 0$, puis se démontre par récurrence sur k).

Par linéarité, il en résulte que si $P \in \mathbb{K}[X]$ est un polynôme quelconque, on aura $P(f_A) = f_{P(A)}$.

Ainsi, si A est diagonalisable, il existe un polynôme scindé à racines simples $P \in \mathbb{K}[X]$ tel que $P(A) = 0$. On en déduit $P(f_A) = f_{0_n} = 0$, donc P est aussi un polynôme annulateur scindé à racines simples de f_A, et f_A est donc diagonalisable.

La réciproque se démontre exactement de la même façon, puisque, dire que, si $B \in \mathcal{M}_n(\mathbb{K})$ est telle que f_B est l'application nulle, cela signifie que la matrice B est nulle.

• *Utilisation des sous-espaces propres*

Nous allons ici caractériser les sous-espaces propres de f_A en fonction de ceux de f_A.

Soit $\lambda \in \mathbb{K}$. On a :

$$f_A(M) = \lambda M \iff (A - \lambda \mathrm{I}_n)M = 0$$
$$\iff \mathrm{Im}(M) \subset \mathrm{Ker}(A - \lambda \mathrm{I}_n)$$

(en confondant comme d'habitude chaque matrice avec l'application linéaire qui lui est canoniquement associée).

Or l'ensemble des matrices M telles que $\mathrm{Im}(M) \subset \mathrm{Ker}(A - \lambda \mathrm{I}_n)$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$, qui est isomorphe à l'espace vectoriel des applications linéaires de $\mathbb{K}^n$ dans $\mathrm{Ker}(A - \lambda \mathrm{I}_n)$, donc qui est de dimension $n \times \dim(\mathrm{Ker}(A - \lambda \mathrm{I}_n))$.

Ainsi :

$$\dim(\mathrm{Ker}(f_A - \lambda \mathrm{Id}_{\mathcal{M}_n(\mathbb{K})}))$$
$$= n \times \dim(\mathrm{Ker}(A - \lambda \mathrm{I}_n)).$$

Par conséquent, compte tenu de $\mathrm{Sp}(f_A) = \mathrm{Sp}(A)$:

$$\sum_{\lambda \in \mathrm{Sp}(A)} \dim(\mathrm{Ker}(A - \lambda \mathrm{I}_n)) = n = \dim \mathbb{K}^n$$

équivaut à :

$$\sum_{\lambda \in \mathrm{Sp}(f_A)} \dim(\mathrm{Ker}(f_A - \lambda \mathrm{Id}_{\mathcal{M}_n(\mathbb{K})}))$$
$$= n^2 = \dim(\mathcal{M}_n(\mathbb{K}))\,,$$

ce qui montre que A est diagonalisable si et seulement si f_A est diagonalisable.

3) Pour tout $M \in \mathcal{M}_n(\mathbb{K})$:

$$g_A(M) = \lambda M \iff f_{{}^tA}({}^tM) = \lambda\, {}^tM\,.$$

Donc $\mathrm{Sp}(g_A) = \mathrm{Sp}(f_{{}^tA})$ et :

$$M \in E_\lambda(g_A) \iff {}^tM \in E_\lambda(f_{{}^tA})\,,$$

d'où $\dim E_\lambda(g_A) = \dim E_\lambda(f_{{}^tA})$.

Par conséquent, si A est diagonalisable, il en est de même de ${}^tA$ donc de $f_{{}^tA}$ d'après **2)b)**. On aura donc :

$$\sum_{\lambda \in \mathrm{Sp}(f_{{}^tA})} \dim E_\lambda(f_{{}^tA}) = n^2$$

d'où

$$\sum_{\lambda \in \mathrm{Sp}(g_A)} \dim E_\lambda(g_A) = n^2\,,$$

ce qui prouve que g_A est diagonalisable. Ses valeurs propres sont celles de $f_{{}^tA}$ donc celles de ${}^tA$ c'est-à-dire celles de A.

4) On a : $u = f_A + g_B$.

D'après les questions précédentes, f_A et g_B sont diagonalisables. De plus, ces endomorphismes commutent puisque, pour tout $M \in \mathcal{M}_n(\mathbb{K})$:

$$f_A \circ g_B(M) = AMB = g_B \circ f_A(M)\,.$$

D'après l'exemple de **[SF4.7]** page 299, il existe une base de $\mathcal{M}_n(\mathbb{K})$ formée de vecteurs propres communs à f_A et g_B. Ces vecteurs formeront donc aussi une base de vecteurs propres de u, donc u est diagonalisable.

On peut préciser que les valeurs propres de u sont alors les sommes des valeurs propres de f_A et g_B, donc aussi les sommes des valeurs propres de A et de B.

Exercice 4.11

1) Nous allons démontrer l'égalité proposée par récurrence sur k.

- Elle est évidemment vérifiée pour $k = 0,\ 1$.
- Si on la suppose vérifiée à l'ordre k, alors :

$$A^{k+1}B - BA^{k+1}$$
$$= A(\underbrace{A^kB - BA^k}_{=kA^k}) + (\underbrace{AB - BA}_{=A})A^k$$
$$= kA^{k+1} + A^{k+1} = (k+1)A^{k+1},$$

ce qui prouve le résultat à l'ordre $k+1$ et achève la récurrence.

2) L'application $\varphi\colon M \mapsto MB - BM$ est bien un endomorphisme de $\mathcal{M}_n(\mathbb{K})$ (vérification facile).

La question précédente donne :

$$\forall k \in \mathbb{N},\ \varphi(A^k) = kA^k\,.$$

Si A n'était pas nilpotente, on aurait $A^k \neq 0$ pour tout entier k, et l'égalité ci-dessus montre que, dans ce cas, tout entier k serait valeur propre de φ (associée au vecteur propre A^k). Mais cela est impossible car φ, endomorphisme d'un espace vectoriel de dimension finie, ne possède qu'un nombre fini de valeurs propres !

En conclusion, on a prouvé par l'absurde que A est nilpotente.

Exercice 4.12

• Si f admet une valeur propre réelle λ et si x est un vecteur propre associé, alors la droite $\mathbb{R}.x$ répond à la question.

• Dans le cas où $\mathrm{Sp}_{\mathbb{R}}(f) = \emptyset$, nous proposons deux solutions.

– *Première méthode*

Soit $\mathscr{B}$ une base de E, et $A \in \mathcal{M}_n(\mathbb{R})$ la matrice de f dans cette base.

A admet donc au moins deux valeurs propres complexes non réelles conjuguées $\lambda = a + \mathrm{i}b$ et $\overline{\lambda} = a - \mathrm{i}b$ $((a,b) \in \mathbb{R}^2)$.

Notons alors $V \in \mathcal{M}_{n,1}(\mathbb{C}) \setminus \{0\}$ un vecteur propre de A associé à la valeur propre $a+\mathrm{i}b$. On peut écrire $V = X + \mathrm{i}Y$ avec X, $Y \in \mathcal{M}_{n,1}(\mathbb{R})$, et $Y \neq 0$ (sinon $AX = \lambda X$ impliquerait λ réel).

$$\begin{aligned} AV = \lambda V &\iff A(X+\mathrm{i}Y) = (a+\mathrm{i}b)(X+\mathrm{i}Y) \\ &\iff AX + \mathrm{i}AY = aX - bY \\ &\qquad\qquad + \mathrm{i}(bX + aY) \\ &\iff \begin{cases} AX = aX - bY \\ AY = bX + aY \end{cases}. \end{aligned}$$

Notons alors x et y les vecteurs de E dont les coordonnées dans la base $\mathscr{B}$ sont respectivement X et Y. On a :

$$\begin{cases} f(x) = ax - by \\ f(y) = bx + ay \end{cases}$$

et donc $\mathrm{Vect}(x,y)$ est stable par f.

Si la famille (x,y) était liée, alors comme $y \neq 0$ il existerait $\mu \in \mathbb{R}$ tel que $x = \mu y$, d'où $f(y) = (b\mu + a)y$, et y serait vecteur propre de f, ce qui est exclu par hypothèse.

On a donc $\dim \mathrm{Vect}(x,y) = 2$, et on a trouvé un plan stable par f.

– *Deuxième méthode*

Le polynôme caractéristique χ_f ne possédant pas de racines réelles peut se décomposer en produits de facteurs irréductibles de $\mathbb{R}[X]$ sous la forme :

$$\chi_f = \prod_{i=1}^{n/2} (X^2 + a_i X + b_i)\,,$$

avec $a_i^2 - 4b_i < 0$ (notons que n est forcément pair, puisque tout polynôme à coefficients réels de degré impair admet une racine réelle). D'après le théorème de Cayley-Hamilton :

$$\chi_f(f) = \prod_{i=1}^{n/2} (f^2 + a_i f + b_i \mathrm{Id}_E) = 0_{\mathscr{L}(E)}\,,$$

ce qui implique que l'un au moins des endomorphismes $f^2 + a_i f + b_i \mathrm{Id}_E$ n'est pas injectif (sinon, leur composée le serait). Le noyau de cet endomorphisme n'est alors pas réduit à $\{0\}$: il existe $x \in E \setminus \{0_E\}$ tel que $f^2 + a_i f + b_i \mathrm{Id}_E(x) = 0$, soit $f^2(x) = -a_i f(x) - b_i x$.

Le sous-espace vectoriel $\mathrm{Vect}(x,f(x))$ est donc stable par f ; et il s'agit d'un plan, sinon x serait un vecteur propre de f ce qui est exclu.

Exercice 4.13

• Puisque l'on sait que tout polynôme de l'endomorphisme u commute avec u, on a l'inclusion facile $\mathbb{K}[u] \subset \mathscr{C}(u)$.

• Soit maintenant $v \in \mathscr{C}(u)$. On veut montrer que $v \in \mathbb{K}[u]$, c'est-à-dire qu'il existe un polynôme $P \in \mathbb{K}[X]$ tel que $v = P(u)$.

Par hypothèse, u possède n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$ et E est de dimension n, donc u est diagonalisable ; de plus, chaque sous-espace propre $E_{\lambda_i}(u)$ est de dimension 1. Notons e_i un vecteur non nul de $E_{\lambda_i}(u)$, de sorte que $\mathscr{B} = (e_1, \ldots, e_n)$ est une base de vecteurs propres de u.

Puisque v commute avec u, on sait (cf. **[S4.6]**) que chaque $E_{\lambda_i}(u)$ est stable par v ; on en déduit que le vecteur $v(e_i)$ est colinéaire à e_i, donc il existe un scalaire μ_i tel que $v(e_i) = \mu_i e_i$: la base $\mathscr{B}$ est donc aussi une base de vecteurs propres de v, pour les valeurs propres μ_i.

Dans cette base $\mathscr{B}$, les matrices de u et v sont diagonales :

$$M_{\mathscr{B}}(u) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

et

$$\text{et } M_{\mathscr{B}}(v) = \mathrm{diag}(\mu_1, \ldots, \mu_n)\,.$$

Puisque pour tout entier k,

$$M_{\mathscr{B}}(u^k) = \mathrm{diag}(\lambda_1^k, \ldots, \lambda_n^k)\,,$$

l'existence de $P \in \mathbb{K}[X]$ tel que $v = P(u)$ équivaut donc à celle de P tel que $\mu_i = P(\lambda_i)$ pour tout i.

Or un tel polynôme P existe puisque les λ_i sont distincts, d'après les résultats du cours sur les polynômes d'interpolation de Lagrange (voir **[S1.12]**).

Il existe donc bien un polynôme P tel que $v = P(u)$, donc $v \in \mathbb{K}[u]$, ce qui prouve l'inclusion cherchée.

Exercice 4.14

1) a) Le polynôme caractéristique χ_A de A est scindé dans $\mathbb{C}$; il peut donc s'écrire sous la forme : $\chi_A(X) = \prod_{k=1}^{n}(X - \lambda_k)$, où les λ_k sont les valeurs propres, distinctes ou non, de A.

On en déduit : $\chi_A(B) = \prod_{k=1}^{n}(B - \lambda_k \mathrm{I}_n)$.

Or les matrices $B - \lambda_k \mathrm{I}_n$ sont inversibles puisque les λ_k ne sont pas valeurs propres de B par hypothèse. Leur produit est donc aussi une matrice inversible, donc :

$$\chi_A(B) \text{ est inversible.}$$

b) L'implication $X = 0 \implies AX = XB$ est immédiate.

Supposons maintenant $AX = XB$. Alors :

$$\begin{aligned} A^2X &= A(AX) = A(XB) \\ &= (AX)B = (XB)B = XB^2, \end{aligned}$$

et, par une récurrence simple, on en déduit $A^kX = XB^k$ pour tout $k \in \mathbb{N}$.

Si $P = \sum_{k=0}^{p} a_k x^k \in \mathbb{C}[x]$ est un polynôme, on a donc :

$$\begin{aligned} P(A)X &= \sum_{k=0}^{p} a_k A^k X \\ &= \sum_{k=0}^{p} a_k X B^k = XP(B)\,. \end{aligned}$$

En prenant en particulier pour P le polynôme caractéristique de A, on a $\chi_A(A)X = X\chi_A(B)$. Or, d'après le théorème de Cayley-Hamilton, $\chi_A(A) = 0$, donc $X\chi_A(B) = 0$. La matrice $\chi_A(B)$ étant inversible d'après la question **1)a)**, on en déduit que $X = 0$.

On a donc bien démontré que, lorsque A et B ont des valeurs propres distinctes :

$$AX = XB \iff X = 0\,.$$

c) Considérons l'application

$$f : \begin{cases} \mathcal{M}_n(\mathbb{C}) & \longrightarrow & \mathcal{M}_n(\mathbb{C}) \\ X & \longmapsto & AX - XB\,. \end{cases}$$

Cette application est un endomorphisme de $\mathcal{M}_n(\mathbb{C})$: en effet, elle est bien à valeurs dans $\mathcal{M}_n(\mathbb{C})$, et la linéarité $\big(f(\lambda X_1 + X_2) = \lambda f(X_1) + f(X_2)\big)$ ne pose pas de problème.

On a en fait démontré à la question précédente que : $f(X) = 0 \iff X = 0$. Autrement dit, f est injective. Comme il s'agit d'un endomorphisme d'un espace vectoriel de dimension finie, elle est donc bijective, ce qui se traduit par :

$$\begin{gathered} \forall\, M \in \mathcal{M}_n(\mathbb{C})\,, \\ \exists ! X \in \mathcal{M}_n(\mathbb{C}) \text{ tel que } M = f(X) \end{gathered}$$

ce qui est exactement le résultat demandé.

2) On sait que le polynôme caractéristique de B est égal au polynôme caractéristique de ${}^tB$, donc B et ${}^tB$ ont mêmes valeurs propres.

Soit donc λ une valeur propre commune à A et B.

Puisque λ est valeur propre de A, il existe une matrice colonne $U = {}^t\begin{pmatrix} u_1 & \ldots & u_n \end{pmatrix}$ dans $\mathcal{M}_{n,1}(\mathbb{C})$, non nulle, telle que $AU = \lambda U$.

Puisque λ est valeur propre de ${}^tB$, il existe une matrice colonne $V = {}^t\begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$ dans $\mathcal{M}_{n,1}(\mathbb{C})$, non nulle, telle que ${}^tBV = \lambda V$.

On considère alors la matrice carrée d'ordre n : $X = U{}^tV$. Le terme d'indice (i,j) de cette matrice est $u_i v_j$. U et V étant non nuls, il existe i_0 tel que $u_{i_0} \neq 0$ et j_0 tel que $v_{j_0} \neq 0$. Le terme d'indice (i_0,j_0) de la matrice X n'est pas nul, la matrice X n'est donc pas la matrice nulle, et l'on a :

$$AX = AU{}^tV = \lambda U{}^tV = \lambda X$$

et

$$XB = U\,{}^{\mathrm{t}}VB = U\,{}^{\mathrm{t}}({}^{\mathrm{t}}BV)$$
$$= U\,{}^{\mathrm{t}}(\lambda V) = \lambda U\,{}^{\mathrm{t}}V = \lambda X\,.$$

Ainsi, si A et B ont une valeur propre commune, il existe une matrice X non nulle telle que $AX = XB$.

On peut résumer les questions **1)** et **2)** de l'exercice par :

$$\mathrm{Sp}(A) \cap \mathrm{Sp}(B) \neq \emptyset \iff$$
$$\exists\, X \in \mathcal{M}_n(\mathbb{C}) \setminus \{0\} \text{ tel que } AX = XB\,.$$

3) Soit donc $X \in \mathcal{M}_n(\mathbb{C})$, non nulle, telle que $AX = XB$, et soit r le rang de X.
D'après le cours (conséquence du théorème de Gauss-Jordan **[S3.13]**), on sait qu'il existe deux matrices inversibles P et Q telles que $X = PJ_rQ$ où J_r est la matrice écrite par blocs

$$J_r = \begin{bmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{bmatrix}.$$

On a alors $APJ_rQ = PJ_rQB$ soit $(P^{-1}AP)J_r = J_r(QBQ^{-1})$.

Notons $A' = P^{-1}AP$ et $B' = QBQ^{-1}$. On écrit ces deux matrices par blocs :

$$A' = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \text{ et } B' = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

avec A_1, $B_1 \in \mathcal{M}_r(\mathbb{C})$.
Alors, en faisant le produit par blocs :

$$A'J_r = \begin{bmatrix} A_1 & 0 \\ A_3 & 0 \end{bmatrix} \text{ et } J_rB' = \begin{bmatrix} B_1 & B_2 \\ 0 & 0 \end{bmatrix}.$$

Donc la relation $A'J_r = J_rB'$ est équivalente à $A_1 = B_1$, $A_3 = 0$ et $B_2 = 0$.

Par conséquent,

$$A' = \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix} \text{ et } B' = \begin{bmatrix} A_1 & 0 \\ B_3 & B_4 \end{bmatrix}.$$

On calcule alors les polynômes caractéristiques à l'aide d'un déterminant par blocs :

$$\chi_{A'} = \begin{vmatrix} XI_r - A_1 & -A_2 \\ 0 & XI_{n-r} - A_4 \end{vmatrix}$$
$$= \chi_{A_1}.\chi_{A_4}$$

et de même : $\chi_{B'} = \chi_{A_1}.\chi_{B_4}$.

De plus, $\chi_A = \chi_{A'}$ puisque A et A' sont semblables, et, de même, $\chi_B = \chi_{B'}$.
Donc, d'après le calcul précédent, χ_{A_1} divise χ_A et χ_B. Les r racines complexes (distinctes ou non) de χ_{A_1} sont donc aussi racines de χ_A et de χ_B ; ce sont donc r valeurs propres communes à A et à B.

Exercice 5.1

• On munit $\mathbb{R}^n$ du produit scalaire canonique.
On applique l'inégalité de Cauchy-Schwarz aux vecteurs $u = \left(\sqrt{x_1}, \ldots, \sqrt{x_n}\right)$ et $v = \left(\dfrac{1}{\sqrt{x_1}}, \ldots, \dfrac{1}{\sqrt{x_n}}\right)$ (cela est possible car les x_i sont strictement positifs).
On a donc $\langle u \,|\, v\rangle^2 \leqslant \|u\|^2 \, \|v\|^2$, soit :

$$n^2 \leqslant \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \frac{1}{x_i}\right).$$

Compte tenu de l'hypothèse $\sum_{i=1}^{n} x_i = 1$, on obtient bien l'inégalité :

$$\sum_{i=1}^{n} \frac{1}{x_i} \geqslant n^2\,.$$

• Il y a égalité si et seulement si les vecteurs u et v sont colinéaires, ce qui se traduit par l'existence d'un réel λ tel que :

$$\forall\, i \in [\![1\,;n]\!],\ \frac{1}{\sqrt{x_i}} = \lambda\sqrt{x_i}\,,$$

soit $x_i = 1/\lambda$ pour tout i, et puisque $\sum_{i=1}^{n} x_i = 1$, cela équivaut à :

$$\forall\, i \in [\![1\,;n]\!],\ x_i = \frac{1}{n}\,.$$

Exercice 5.2

1) L'application :

$$\mathrm{tr}: \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ A & \longmapsto & \mathrm{tr}(A) \end{cases}$$

est une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$, directement d'après le cours.

Cette forme linéaire étant non nulle, son noyau est un hyperplan H de $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ de dimension $\dim(\mathcal{M}_n(\mathbb{R})) - 1 = n^2 - 1$.

2) Notons p la projection orthogonale sur H, et q la projection orthogonale sur la droite vectorielle $H^{\perp}$, de sorte que $p + q = \mathrm{Id}_{\mathcal{M}_n(\mathbb{R})}$.

La distance d de J à H est $d = \|J - p(J)\| = \|q(J)\|$. Il nous suffit donc de déterminer $q(J)$.

> ✓ On utilise le fait qu'il est bien plus facile de déterminer la projection orthogonale sur une droite vectorielle que sur un hyperplan.

Pour déterminer $q(J)$, il nous suffit de déterminer un vecteur non nul orthogonal à H. Or, par définition du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$, pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, $\operatorname{tr} A = \langle A \,|\, \mathrm{I}_n \rangle$ où I_n est la matrice identité d'ordre n. Donc

$$A \in H \iff \langle A \,|\, \mathrm{I}_n \rangle = 0\,,$$

et par conséquent, $H^\perp = \mathbb{R}.\mathrm{I}_n$.

Il ne reste plus qu'à appliquer la formule du projeté orthogonal sur une droite vectorielle :

$$q(J) = \frac{\langle J \,|\, \mathrm{I}_n \rangle}{\|\mathrm{I}_n\|^2} \mathrm{I}_n\,,$$

d'où :

$$d = \|q(J)\| = \frac{|\langle J \,|\, \mathrm{I}_n \rangle|}{\|\mathrm{I}_n\|}\,.$$

On calcule :

$$\langle J \,|\, \mathrm{I}_n \rangle = \operatorname{tr} J = n$$

et

$$\|\mathrm{I}_n\|^2 = \operatorname{tr}(\mathrm{I}_n) = n \text{ d'où } \|\mathrm{I}_n\| = \sqrt{n}\,.$$

Conclusion : la distance de J à H est égale à $\sqrt{n}$.

Exercice 5.3

On va naturellement se placer ici dans $\mathbb{R}^n$ muni de sa structure euclidienne canonique.

• Notons $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ les vecteurs de l'énoncé (fixés). Notons aussi z le vecteur $z = (1,1,\ldots,1)$, de sorte que les $y_i - ax_i - b$ sont les coordonnées du vecteur $y - ax - bz$.

Ainsi $\varphi(a,b) = \|y - (ax + bz)\|^2$.

Lorsque a et b décrivent $\mathbb{R}$, le vecteur $ax + bz$ décrit le sous-espace vectoriel de $\mathbb{R}^n$: $P = \operatorname{Vect}(x,z)$, qui est un plan : en effet, x et z ne sont pas colinéaires puisque les x_i ne sont pas tous égaux.

Donc $\varphi(a,b)$ est le carré de la distance entre y et un vecteur de P.

• D'après le cours, directement, cette distance admet un minimum absolu lorsque le vecteur $ax + bz$ décrit P ; ce minimum est la distance de y à P, et il est atteint pour un unique vecteur de P, à savoir le projeté orthogonal de y sur P.

• Il ne reste plus qu'à déterminer ce projeté orthogonal ; pour cela, il y a deux façons de faire : soit utiliser le cours sur les espaces euclidiens et les formules correspondantes, soit utiliser la théorie des fonctions de plusieurs variables.

1) *Utilisation de la projection orthogonale, première façon.*

Nous allons appliquer ici la formule donnant directement le projeté orthogonal ; pour cela, il suffit de trouver une base orthogonale (e_1,e_2) du plan $P = \operatorname{Vect}(x,z)$. Pour ce faire, on peut utiliser le procédé d'orthonormalisation de Schmidt ; ce procédé est pénible en général, mais ici le nombre des calculs est réduit.

En effet, si on choisit $e_1 = z$, la formule nous donne directement :

$$e_2 = x - \frac{\langle e_1 \,|\, x \rangle}{\|e_1\|^2}\, e_1\,,$$

et puisque

$$\langle e_1 \,|\, x \rangle = \langle z \,|\, x \rangle = \sum_{i=1}^{n} x_i = n\,\overline{x}$$

et que

$$\|e_1\|^2 = \|z\|^2 = n\,,$$

on a $e_2 = x - \overline{x}\, e_1$ (puisque les x_i ne sont pas tous égaux, e_2 n'est pas nul).

Le projeté orthogonal de x sur P est alors donné par :

$$\begin{aligned} p_P(y) &= \frac{\langle e_1 \,|\, y \rangle}{\|e_1\|^2}\, e_1 + \frac{\langle e_2 \,|\, y \rangle}{\|e_2\|^2}\, e_2 \\ &= \overline{y}\, e_1 + \frac{\langle x \,|\, y \rangle - n\overline{x}\,\overline{y}}{\|e_2\|^2} \big(x - \overline{x}\, e_1\big)\,. \end{aligned}$$

$$\begin{aligned} \text{Or : } \|e_2\|^2 &= \sum_{i=1}^{n} (x_i - \overline{x})^2 \\ &= \|x\|^2 - 2\overline{x} \sum_{i=1}^{n} x_i + n\overline{x}^{\,2} \\ &= \|x\|^2 - n\overline{x}^{\,2} \end{aligned}$$

donc le projeté orthogonal s'écrit $p_P(y) = ax + bz$ avec :

$$a = \frac{\langle x \,|\, y \rangle - n\overline{x}\,\overline{y}}{\|x\|^2 - n\overline{x}^{\,2}} \quad \text{et} \quad b = \overline{y} - a\overline{x}\,.$$

2) *Utilisation de la projection orthogonale, seconde façon.*

Il est en général préférable d'éviter le procédé d'orthonormalisation pour trouver le projeté orthogonal. Pour cela, il suffit de revenir à la définition d'une projection.

Si l'on note $p_P(y) = ax + bz$ le projeté orthogonal de y sur $\mathrm{Vect}(x,z)$, on a $y - p_P(y) \in \mathrm{Vect}(x,z)^\perp$, ce qui se traduit par :

$$\langle y - p_P(y) \,|\, x\rangle = \langle y - p_P(y) \,|\, z\rangle = 0\,,$$

d'où le système :

$$\begin{cases} a\,\|x\|^2 + b\,\langle z\,|\,x\rangle = \langle y\,|\,x\rangle \\ a\,\langle x\,|\,z\rangle + b\,\|z\|^2 = \langle y\,|\,z\rangle \end{cases}.$$

Puisque $\langle x\,|\,z\rangle = n\overline{x}$, $\langle y\,|\,z\rangle = n\overline{y}$ et $\|z\|^2 = n$, ce système de 2 équations à 2 inconnues se résout aisément, et permet d'obtenir de façon plus rapide les valeurs de a et de b.

3) *Utilisation des fonctions de plusieurs variables.*

On a démontré que la fonction φ admet un minimum absolu sur $\mathbb{R}^2$, et que ce minimum est atteint pour une unique valeur du couple (a,b). La fonction φ étant évidemment de classe $\mathscr{C}^1$, ce couple (a,b) correspond nécessairement à un point critique (le seul) de φ.

a et b sont donc *les* solutions du système :

$$\frac{\partial\varphi}{\partial a}(a,b) = \frac{\partial\varphi}{\partial b}(a,b) = 0\,,$$

soit :

$$\begin{cases} -\displaystyle\sum_{i=1}^{n} x_i(y_i - ax_i - b) = 0 \\ -\displaystyle\sum_{i=1}^{n} (y_i - ax_i - b) = 0\,. \end{cases}$$

La deuxième équation de ce système donne immédiatement la relation :

$$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} x_i + nb$$

d'où $\overline{y} = a\overline{x} + b$.

La première s'écrit alors :

$$a\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} x_i$$

d'où

$$a\,\|x\|^2 = \langle x\,|\,y\rangle - (\overline{y} - a\overline{x})n\overline{x}\,,$$

ce qui permet de retrouver la valeur de a obtenue précédemment.

✓ Cette méthode semble ici plus rapide ; cependant, il a fallu déjà montrer que le point critique obtenu correspondait bien à un minimum absolu, donc utiliser les résultats théoriques sur les projections orthogonales (mais pas les formules).

✓ Le problème abordé par cet exercice est celui dit « des moindres carrés » : il s'agit de trouver la droite d'équation $y = ax+b$ qui approche « le mieux » le nuage de points (x_i,y_i). Cette droite s'appelle la droite de régression.

Exercice 5.4

1) • L'application $y \mapsto \|x - y\|$ étant continue sur le compact C, elle y est bornée et atteint ses bornes. Il existe donc un $x_0 \in C$ tel que

$$\|x - x_0\| = \inf_{y\in C} \|x - y\| = d(x,C).$$

• Supposons qu'il existe deux points x_0 et x_1 de C qui réalisent ce minimum. Puisque C est convexe, le milieu $y = \dfrac{x_0 + x_1}{2}$ du segment $[x_0\,;x_1]$ appartient à C.
L'identité du parallélogramme :

$$2\big(\|a\|^2 + \|b\|^2\big) = \|a+b\|^2 + \|a-b\|^2$$

appliquée aux vecteurs :

$$a = x_0 - x \quad\text{et}\quad b = x_1 - x\,,$$

donne alors :

$$4\,\|y - x\|^2 + \|x_0 - x_1\|^2 = 2\,\|x_0 - x\|^2 + 2\,\|x_1 - x\|^2\,,$$

et puisque :

$$\|x_0 - x\|^2 = \|x_1 - x\|^2 = d(x,C)^2$$

et :

$$\|y - x\|^2 \geqslant d(x,C)^2\,,$$

on en déduit $\|x_0 - x_1\|^2 \leqslant 0$, d'où $x_0 = x_1$.
Cela démontre donc l'unicité de x_0.

2) On conserve les mêmes notations. Soit $y \in C$. Pour tout $t \in\,]0\,;1]$, le vecteur $ty + (1-t)x_0$ appartient à C donc :

$$\|x - ty - (1-t)x_0\|^2 \geqslant \|x - x_0\|^2\,,$$

soit :

$$\|(x - x_0) - t(y - x_0)\|^2 \geqslant \|x - x_0\|^2\,.$$

En développant cette expression, on obtient :

$$t^2 \|y - x_0\|^2 - 2t \langle x - x_0 \,|\, y - x_0 \rangle \geqslant 0 \,.$$

En divisant par $t > 0$ puis en faisant tendre t vers 0, on en déduit :

$$\langle x - x_0 \,|\, y - x_0 \rangle \leqslant 0 \,,$$

ce qu'il fallait démontrer.

Exercice 5.5

1) • Montrons d'abord que E est bien un sous-espace vectoriel de $\mathscr{C}([0\,;+\infty[,\mathbb{R})$.

– Déjà, E est bien inclus dans $\mathscr{C}([0\,;+\infty[,\mathbb{R})$, par définition.

– E est non vide puisqu'il contient la fonction nulle.

– Il est clair que $f \in E \implies \lambda f \in E$ pour tout réel λ.

– Il reste à montrer que E est stable par addition. Pour cela, on utilise l'inégalité bien connue :

$$\forall\,(a,b) \in \mathbb{R}^2,\ |ab| \leqslant \frac{1}{2}(a^2 + b^2)\,,$$

dont la démonstration est élémentaire (identités remarquables).
Si f et g sont dans E on aura donc pour tout $t \geqslant 0$:

$$|f(t)g(t)|\,\mathrm{e}^{-t} \leqslant \frac{1}{2}\left(f^2(t)\mathrm{e}^{-t} + g^2(t)\mathrm{e}^{-t}\right)$$

d'où il résulte, d'après les théorèmes de comparaison pour les fonctions à valeurs positives, que l'intégrale $\displaystyle\int_0^{+\infty} f(t)g(t)\mathrm{e}^{-t}\,\mathrm{d}t$ est absolument convergente donc convergente.
Puisque

$$(f(t)+g(t))^2\mathrm{e}^{-t} = f^2(t)\mathrm{e}^{-t} + g^2(t)\mathrm{e}^{-t} + 2f(t)g(t)\mathrm{e}^{-t}\,,$$

on en déduit la convergence de l'intégrale $\displaystyle\int_0^{+\infty}(f(t)+g(t))^2\mathrm{e}^{-t}\,\mathrm{d}t$, c'est-à-dire que $f+g \in E$.

• Si P est une fonction polynôme, elle est continue sur $[0\,;+\infty[$ et

$$\lim_{t\to+\infty} t^2 P^2(t)\mathrm{e}^{-t} = 0\,,$$

ce qui prouve que

$$P^2(t)\mathrm{e}^{-t} \underset{t\to+\infty}{=} o\left(\frac{1}{t^2}\right).$$

La fonction $f\colon t \mapsto \dfrac{1}{t^2}$ étant intégrable positive au voisinage de $+\infty$, on en déduit qu'il en est de même de $t \mapsto P^2(t)\mathrm{e}^{-t}$ et par conséquent E contient l'ensemble des fonctions polynômes.

• On a vu que, si f et g appartiennent à E, alors l'intégrale $\displaystyle\int_0^{+\infty} f(t)g(t)\mathrm{e}^{-t}\,\mathrm{d}t$ est convergente ; cela prouve que la définition de $\langle f \,|\, g\rangle$ a bien un sens.

La démonstration du fait qu'il s'agit bien d'une forme bilinéaire symétrique définie positive est alors en tout point semblable à celle faite en cours pour le produit scalaire usuel $(f,g) \mapsto \displaystyle\int fg$.

2) • D'après la formule de Leibniz :

$$\begin{aligned}
&\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^n\mathrm{e}^{-x}\right)\\
&= \sum_{k=0}^{n}\binom{n}{k}\frac{\mathrm{d}^k}{\mathrm{d}x^k}(x^n)\frac{\mathrm{d}^{n-k}}{\mathrm{d}x^{n-k}}(\mathrm{e}^{-x})\\
&= \sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}\frac{n!}{(n-k)!}x^{n-k}\mathrm{e}^{-x}\,,
\end{aligned}$$

donc

$$L_n(x) = \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^{n-k}}{(n-k)!}x^{n-k}$$

est bien une fonction polynôme, de degré n et de coefficient dominant $\dfrac{(-1)^n}{n!}$.

• On suppose que $p > q$. On a alors en faisant une intégration par parties :

$$\begin{aligned}
&\langle L_p \,|\, L_q\rangle\\
&= \frac{1}{p!}\int_0^{+\infty}\frac{\mathrm{d}^p}{\mathrm{d}x^p}\left(x^p\mathrm{e}^{-x}\right)L_q(x)\,\mathrm{d}x\\
&= \frac{1}{p!}\left[\frac{\mathrm{d}^{p-1}}{\mathrm{d}x^{p-1}}\left(x^p\mathrm{e}^{-x}\right)L_q(x)\right]_0^{+\infty}\\
&\quad - \frac{1}{p!}\int_0^{+\infty}\frac{\mathrm{d}^{p-1}}{\mathrm{d}x^{p-1}}\left(x^p\mathrm{e}^{-x}\right)L_q'(x)\,\mathrm{d}x\,.
\end{aligned}$$

Or en utilisant toujours la formule de Leibniz, il est facile de voir que $\dfrac{\mathrm{d}^{p-1}}{\mathrm{d}x^{p-1}}\left(x^p\mathrm{e}^{-x}\right)$ s'écrit sous la forme $\mathrm{e}^{-x}P(x)$ avec P polynôme tel que $P(0) = 0$, donc la valeur du terme entre crochets ci-dessus est nulle (ce qui justifie a posteriori l'intégration par

parties) et l'on trouve :

$$\langle L_p \,|\, L_q\rangle = -\frac{1}{p!}\int_0^{+\infty} \frac{\mathrm{d}^{p-1}}{\mathrm{d}x^{p-1}}\left(x^p \mathrm{e}^{-x}\right) L_q'(x)\,\mathrm{d}x\,.$$

En faisant ainsi p intégrations par parties consécutives, on obtient finalement :

$$\langle L_p \,|\, L_q\rangle = \frac{(-1)^p}{p!}\int_0^{+\infty} x^p \mathrm{e}^{-x}\frac{\mathrm{d}^p}{\mathrm{d}x^p}\big(L_q(x)\big)\,\mathrm{d}x.$$

Or L_q est un polynôme de degré q donc si $p > q$ sa dérivée p^{e} est nulle. On obtient donc $\langle L_p \,|\, L_q\rangle = 0$.

Par symétrie on obtient le même résultat si $q > p$, ce qui prouve que la famille $(L_n)_{n\in\mathbb{N}}$ est orthogonale.

• En reprenant le calcul ci-dessus on a :

$$\|L_n\|^2 = \frac{(-1)^n}{n!}\int_0^{+\infty} x^n \mathrm{e}^{-x} L_n^{(n)}(x)\,\mathrm{d}x\,.$$

Puisque L_n est un polynôme de degré n et de terme dominant $\dfrac{(-1)^n}{n!}X^n$, sa dérivée n^{e} est égale à $(-1)^n$. Il reste donc :

$$\|L_n\|^2 = \frac{1}{n!}\int_0^{+\infty} x^n \mathrm{e}^{-x}\,\mathrm{d}x\,.$$

Le résultat $\displaystyle\int_0^{+\infty} x^n \mathrm{e}^{-x}\,\mathrm{d}x = n!$ est connu (il se démontre aisément par récurrence sur n) et l'on obtient donc $\|L_n\| = 1$.

En conclusion, $(L_n)_{n\in\mathbb{N}}$ est une famille orthonormale de polynômes pour le produit scalaire considéré.

✓ Les L_n s'appellent les polynômes de Laguerre.

3) On calcule facilement :

$$\|f_\alpha\|^2 = \langle f_\alpha \,|\, f_\alpha\rangle = \int_0^{+\infty} \mathrm{e}^{(-2\alpha-1)x}\,\mathrm{d}x = \frac{1}{2\alpha+1}\cdot$$

D'autre part,

$$\langle f_\alpha \,|\, L_n\rangle = \frac{1}{n!}\int_0^{+\infty} \mathrm{e}^{-\alpha x}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^n \mathrm{e}^{-x}\right)\mathrm{d}x$$

et en faisant comme précédemment n intégrations par parties successives on obtient :

$$\begin{aligned}\langle f_\alpha \,|\, L_n\rangle &= \frac{(-1)^n}{n!}\int_0^{+\infty}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\mathrm{e}^{-\alpha x}\right)x^n\mathrm{e}^{-x}\,\mathrm{d}x\\ &= \frac{\alpha^n}{n!}\int_0^{+\infty}\mathrm{e}^{-(\alpha+1)x}x^n\,\mathrm{d}x \quad {\scriptstyle (t=(\alpha+1)x)}\\ &= \frac{\alpha^n}{n!}\frac{1}{(\alpha+1)^{n+1}}\int_0^{+\infty} t^n\mathrm{e}^{-t}\,\mathrm{d}t\\ &= \frac{\alpha^n}{(\alpha+1)^{n+1}}\,.\end{aligned}$$

On en déduit :

$$\begin{aligned}\sum_{n=0}^{+\infty}\langle f_\alpha \,|\, L_n\rangle^2 &= \frac{1}{(\alpha+1)^2}\sum_{n=0}^{+\infty}\left(\frac{\alpha}{\alpha+1}\right)^{2n}\\ &= \frac{1}{(\alpha+1)^2}\frac{1}{1-\left(\frac{\alpha}{\alpha+1}\right)^2}\\ &= \frac{1}{2\alpha+1} = \|f_\alpha\|^2\,.\end{aligned}$$

Bien sûr, ce résultat serait immédiat si l'on savait que la famille $(L_n)_{n\in\mathbb{N}}$ est totale dans E (c'est l'égalité de Parseval), mais cela n'est pas très facile à démontrer (le résultat du cours sur les polynômes orthogonaux a été établi lorsqu'on considère des fonctions continues sur un *segment*, puisque la démonstration de ce résultat utilise le théorème de Weierstrass, lui-même admis). On peut cependant démontrer que f_α appartient à l'adhérence de $\mathrm{Vect}\big((L_n)_{n\in\mathbb{N}}\big)$. En effet, si l'on note p_n la projection orthogonale sur $\mathrm{Vect}(L_0,\ldots,L_n)$, on a

$$\begin{aligned}\|p_n(f_\alpha) - f_\alpha\|^2 &= \|f_\alpha\|^2 - \|p_n(f_\alpha\|^2\\ &= \sum_{k=0}^{+\infty}\langle f_\alpha \,|\, L_k\rangle^2 - \sum_{k=0}^{n}\langle f_\alpha \,|\, L_k\rangle^2\\ &= \sum_{k=n+1}^{+\infty}\langle f_\alpha \,|\, L_k\rangle^2\,,\end{aligned}$$

donc $\displaystyle\lim_{n\to+\infty}\|p_n(f_\alpha) - f_\alpha\| = 0$ (reste d'une série convergente), ce qui signifie que la suite $\big(p_n(f_\alpha)\big)$ converge vers f_α dans l'espace vectoriel normé $(E, \|\cdot\|)$.

Exercice 5.6

On utilisera ici les résultats d'un exemple vu en **[SF5.6]** page 339, à savoir :

$$\mathrm{Ker}\,{}^{\mathrm{t}}AA = \mathrm{Ker}\,A \text{ et } \mathrm{Im}\,{}^{\mathrm{t}}AA = \mathrm{Im}\,{}^{\mathrm{t}}A\,,$$

et d'un exemple vu en **[SF5.9]** page 341, à savoir :

$$(\operatorname{Ker} A)^{\perp} = \operatorname{Im}{}^{t}A\,.$$

1) Puisque

$$\operatorname{Ker} A = \operatorname{Ker}{}^{t}AA = \operatorname{Ker} A^2\,,$$

on a d'après le théorème du rang $\operatorname{rg} A = \operatorname{rg} A^2$ et puisque l'on a toujours l'inclusion $\operatorname{Im} A^2 \subset \operatorname{Im} A$, on en déduit : $\operatorname{Im} A = \operatorname{Im} A^2$. Donc :

$$\operatorname{Im} A = \operatorname{Im}{}^{t}AA = \operatorname{Im}{}^{t}A = (\operatorname{Ker} A)^{\perp}\,,$$

ce qui est le résultat demandé.

2) Dans une base orthonormale adaptée à la décomposition en somme directe $\operatorname{Ker} A \overset{\perp}{\oplus} \operatorname{Im} A$, la matrice de l'endomorphisme a canoniquement associé à A est de la forme :

$$A' = \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix}$$

où B est la matrice de l'endomorphisme induit par a sur $\operatorname{Im} A$. Il existe donc une matrice orthogonale P telle que :

$$A = P \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix} {}^{t}P\,,$$

et de l'égalité ${}^{t}AA = A^2$ on tire ${}^{t}BB = B^2$.

Mais d'après le théorème d'isomorphisme **[S2.17]**, l'endomorphisme induit par a sur $\operatorname{Im} A$ (qui est ici un supplémentaire de $\operatorname{Ker} A$) est un isomorphisme de $\operatorname{Im} A$ sur $\operatorname{Im} A$, c'est-à-dire que la matrice B est inversible.

Donc en multipliant l'égalité ${}^{t}BB = B^2$ par B^{-1} à droite, on obtient ${}^{t}B = B$.

Par conséquent, B est symétrique donc $A = PA'{}^{t}P$ l'est aussi.

Exercice 5.7

1) Notons $\mathscr{B} = (e_1,\dots,e_n)$.
Le terme d'indice (i,j) de la matrice A est la cordonnée de v_j sur e_i, c'est-à-dire $\langle e_i \,|\, v_j\rangle$. Par la formule du produit matriciel :

$$({}^{t}AA)_{ij} = \sum_{k=1}^{n} \langle e_k \,|\, v_i\rangle \langle e_k \,|\, v_j\rangle = \langle v_i \,|\, v_j\rangle$$

donc ${}^{t}AA = G(v_1,\dots,v_n)$.

2) On en déduit :

$$\det G(v_1,\dots,v_n) = \det({}^{t}AA) = (\det A)^2$$

donc

$$\det G(v_1,\dots,v_n) \geqslant 0$$

puis

$$\det G(v_1,\dots,v_n) \neq 0 \iff \det A \neq 0\,,$$

ce qui équivaut à $(v_1,\dots,v_n)$ base de F c'est-à-dire à $(v_1,\dots,v_n)$ libre.

3) On développe $\det G(x,v_1,\dots,v_n)$ en écrivant :

$$x = (x - p(x)) + p(x)$$

puis en utilisant la linéarité du déterminant par rapport à la première colonne :

$$\begin{aligned}
&\det G(x,v_1,\dots,v_n) \\
&= \begin{vmatrix}
\langle x - p(x) \,|\, x\rangle & \langle x \,|\, v_1\rangle & \dots & \langle x \,|\, v_n\rangle \\
\langle x - p(x) \,|\, v_1\rangle & \langle v_1 \,|\, v_1\rangle & \dots & \langle v_1 \,|\, v_n\rangle \\
\vdots & \vdots & & \vdots \\
\langle x - p(x) \,|\, v_n\rangle & \langle v_1 \,|\, v_n\rangle & \dots & \langle v_n \,|\, v_n\rangle
\end{vmatrix} \\
&\quad + \begin{vmatrix}
\langle p(x) \,|\, x\rangle & \langle x \,|\, v_1\rangle & \dots & \langle x \,|\, v_n\rangle \\
\langle p(x) \,|\, v_1\rangle & \langle v_1 \,|\, v_1\rangle & \dots & \langle v_1 \,|\, v_n\rangle \\
\vdots & \vdots & & \vdots \\
\langle p(x) \,|\, v_n\rangle & \langle v_1 \,|\, v_n\rangle & \dots & \langle v_n \,|\, v_n\rangle
\end{vmatrix}\,.
\end{aligned}$$

Or $p(x)$ appartient à $\operatorname{Vect}(v_1,\dots,v_n)$ donc il existe des réels λ_j tels que :

$$p(x) = \sum_{j=1}^{n} \lambda_j v_j\,,$$

ce qui implique que les colonnes $C_0,C_1,\dots,C_n$ du second déterminant ci-dessus sont liées par la relation :

$$C_0 = \sum_{j=1}^{n} \lambda_j C_j$$

donc ce déterminant est nul.

De plus, dans le premier déterminant ci-dessus, les produits scalaires $\langle x - p(x) \,|\, v_i\rangle$ sont nuls puisque $x - p(x)$ est orthogonal à $\operatorname{Vect}(v_1,\dots,v_n)$, donc en développant ce déterminant selon la 1re colonne, on obtient qu'il est égal à :

$$\langle x - p(x) \,|\, x\rangle\, G(v_1,\dots,v_n)\,,$$

ce qui donne la relation demandée.

4) Dans cette question, la famille $(v_1,\dots,v_n)$ est libre, donc $F = \operatorname{Vect}(v_1,\dots,v_n)$.

Puisque $\langle x-p(x)\,|\,x\rangle = \|x-p(x)\|^2$, la relation précédente s'écrit :

$$\forall\, x\in E,\ \det G(x,v_1,\dots,v_n) = \|x-p(x)\|^2\, G(v_1,\dots,v_n)$$

et puisque $\|x-p(x)\| = d(x,F)$ et que $\det G(v_1,\dots,v_n)\neq 0$, on obtient la relation voulue.

Exercice 5.8

• Puisque chaque p_n est une application linéaire, on a, pour tout réel λ et tous vecteurs x, y de E :

$$p_n(\lambda x+y) = \lambda p_n(x) + p_n(y)\,.$$

Les propriétés des limites dans un espace vectoriel normé permettent alors d'en déduire, en faisant tendre n vers $+\infty$:

$$u(\lambda x+y) = \lambda u(x) + u(y)\,,$$

donc u est linéaire, et c'est un endomorphisme de E.

• Pour tout $x\in E$, l'inégalité triangulaire permet d'écrire :

$$\begin{aligned}\big\|p_n^2(x)-u^2(x)\big\| &= \big\|\big(p_n^2(x)-p_n(u(x))\big) \\ &\qquad +\big(p_n(u(x))-u^2(x)\big)\big\| \\ &\leqslant \big\|p_n\big(p_n(x)-u(x)\big)\big\| \\ &\qquad +\big\|p_n(u(x))-u^2(x)\big\| \\ &\leqslant \|p_n(x)-u(x)\| \\ &\qquad +\big\|p_n(u(x))-u^2(x)\big\|\,,\end{aligned}$$

en ayant utilisé aussi la propriété :

$$\forall\, X\in E,\ \|p_n(X)\|\leqslant\|X\|\ ,$$

puisque chaque p_n est un projecteur orthogonal.

Or : $\lim\limits_{n\to+\infty} p_n(x) = u(x)$

et : $\lim\limits_{n\to+\infty} p_n(u(x)) = u(u(x)) = u^2(x)\,,$

donc

$$\lim_{n\to+\infty} p_n^2(x) = u^2(x)\,,$$

et puisque $p_n^2 = p_n$ on en déduit $u^2(x) = u(x)$ c'est-à-dire que u est un projecteur.

• On a enfin :

$$\forall\, x\in E,\ \|p_n(x)\|\leqslant\|x\|\ ,$$

et la continuité de la norme permet alors d'en déduire, quand n tend vers $+\infty$:

$$\forall\, x\in E,\ \|u(x)\|\leqslant\|x\|\ .$$

• La caractérisation des projecteurs orthogonaux permet alors de conclure que u est un projecteur orthogonal.

Exercice 5.9

• *Analyse*

S'il existe une symétrie orthogonale s par rapport à un hyperplan H telle que $s(a)=b$, alors le vecteur $a-s(a)=a-b$ appartient à $H^\perp$.

Puisque $H^\perp$ est une droite et que le vecteur $a-b$ n'est pas nul par hypothèse, c'en est un vecteur de base. Ainsi, H est nécessairement l'hyperplan $(a-b)^\perp$, ce qui montre que s, si elle convient, est unique.

• *Synthèse*

Considérons donc la réflexion par rapport à l'hyperplan $(a-b)^\perp$. Puisque a et b ont mêmes normes on a :

$$\langle a-b\,|\,a+b\rangle = \|a\|^2-\|b\|^2 = 0\,,$$

c'est-à-dire que $a-b\perp a+b$.

On peut donc écrire :

$$a = \underbrace{\frac{a+b}{2}}_{\in H} + \underbrace{\frac{a-b}{2}}_{\in H^\perp}\,,$$

et, par définition d'une symétrie,

$$s(a) = \frac{a+b}{2} - \frac{a-b}{2} = b\,,$$

ce qui montre que s convient.

Exercice 6.1

On commence bien sûr par remarquer que la matrice ${}^tLL$ est symétrique réelle, donc diagonalisable.

De plus, si l'on note :

$$L = \begin{pmatrix} a_1 & a_2 & \dots & a_n\end{pmatrix}\,,$$

cette matrice s'écrit, par blocs,

$${}^tLL = \begin{bmatrix} a_1\,{}^tL & a_2\,{}^tL & \dots & a_n\,{}^tL\end{bmatrix} \quad (*)$$

ce qui montre, puisque $L\neq 0$, que son rang est égal à 1, donc que son noyau est de dimension $n-1$ d'après le théorème du rang.

Puisqu'elle est diagonalisable, 0 en est valeur propre d'ordre $n-1$; et puisque $\operatorname{tr}{}^tLL = \sum_{i=1}^{n} a_i^2 = \|L\|^2$, sa dernière valeur propre est égale à $\|L\|^2$.

D'après $(*)$, son image est la droite vectorielle engendrée par le vecteur colonne ${}^tL$; il s'agit aussi du sous-espace propre associé à la valeur propre $\|L\|^2$, donc l'autre sous-espace propre c'est-à-dire le noyau de ${}^tLL$, est l'hyperplan orthogonal à ${}^tL$, soit l'ensemble des $X \in \mathcal{M}_{n,1}(\mathbb{R})$ tels que $LX = 0$.

✓ Ce dernier résultat pouvait aussi s'obtenir en écrivant le système ${}^tLLX = 0$, mais puisque la matrice ${}^tLL$ est symétrique réelle on sait que ses sous-espaces propres sont orthogonaux deux à deux ; il faut savoir exploiter cette propriété.

✓ Notez que nous avons encore utilisé la propriété importante : les sous-espaces propres d'un endomorphisme associés à une valeur propre non nulle sont inclus dans l'image.

En conclusion, la matrice dans une base orthonormale formée de vecteurs propres de l'endomorphisme associé à ${}^tLL$ est égale à :

$$\operatorname{diag}(0,\dots,0,\|L\|^2)\,,$$

ce qui montre qu'il s'agit de la composée de la projection orthogonale sur la droite de base ${}^tL$ et de l'homothétie de rapport $\|L\|^2$.

Exercice 6.2

Soit $\mathscr{B} = (e_1,\dots,e_n)$ une base orthonormale de E.
Pour tous i, $j \in [\![1\,;n]\!]$ on a :

$$\begin{aligned}\langle e_i - e_j \,|\, e_i + e_j\rangle &= \|e_i\|^2 - \|e_j\|^2\\ &= 1-1 = 0\,,\end{aligned}$$

donc les vecteurs $e_i - e_j$ et $e_i + e_j$ sont orthogonaux.

D'après l'hypothèse il en sera donc de même de leurs images par u c'est-à-dire des vecteurs $u(e_i)-u(e_j)$ et $u(e_i)+u(e_j)$; cela s'écrit :

$$\langle u(e_i) - u(e_j) \,|\, u(e_i) + u(e_j)\rangle = 0$$

d'où $\|u(e_i)\|^2 = \|u(e_j)\|^2$.

Notons alors $\alpha = \|u(e_i)\|$ pour tout i ; on a $\alpha > 0$ car u est bijective donc la famille $\big(u(e_i)\big)$ est une base de E.

Si $i \neq j$, $e_i \perp e_j$ donc on aura aussi $u(e_i) \perp u(e_j)$ et finalement :

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ \langle u(e_i)\,|\,u(e_j)\rangle = \alpha^2\delta_{ij}\,.$$

Si x est un vecteur quelconque de E, il s'écrit dans la base $\mathscr{B}$: $x = \sum_{i=1}^{n} x_i e_i$, donc :

$$\begin{aligned}\|u(x)\|^2 &= \left\|\sum_{i=1}^{n} x_i u(e_i)\right\|^2\\ &= \left\langle \sum_{i=1}^{n} x_i u(e_i) \,\middle|\, \sum_{j=1}^{n} x_j u(e_j)\right\rangle\\ &= \sum_{1\leqslant i,j\leqslant n} x_i x_j \,\langle u(e_i)\,|\,u(e_j)\rangle\\ &= \alpha^2 \sum_{1\leqslant i,j\leqslant n} x_i x_j \delta_{ij}\\ &= \alpha^2 \sum_{i=1}^{n} x_i^2 = \alpha^2\,\|x\|^2\,,\end{aligned}$$

ce qui démontre le résultat cherché.

Exercice 6.3

• Déjà, il est facile de vérifier que u est bien une application linéaire, et qu'elle est à valeurs dans $\mathbb{R}_n[X]$ (si $\deg P \leqslant n$, alors $\deg u(P) \leqslant n$). C'est donc bien un endomorphisme de $\mathbb{R}_n[X]$.

• Soient P, Q dans $\mathbb{R}_n[X]$. Alors :

$$\begin{aligned}&\langle u(P)\,|\,Q\rangle\\ &= \int_{-1}^{1} \big[(1-t^2)P''(t) - 2tP'(t)\big]Q(t)\,\mathrm{d}t\,.\end{aligned}$$

On remarque alors que :

$$(1-t^2)P''(t) - 2tP'(t)$$

est la dérivée de :

$$(1-t^2)P'(t)\,,$$

ce qui suggère de faire une intégrations par parties :

$$\begin{aligned}\langle u(P)\,|\,Q\rangle &= \underbrace{\big[(1-t^2)P'(t)Q(t)\big]_{-1}^{1}}_{=0}\\ &\quad - \int_{-1}^{1} (1-t^2)P'(t)Q'(t)\,\mathrm{d}t\,.\end{aligned}$$

Ainsi,

$$\langle u(P)\,|\,Q\rangle = -\int_{-1}^{1} (1-t^2)P'(t)Q'(t)\,\mathrm{d}t\,,$$

et sous cette forme il est clair que :

$$\langle u(P)\,|\,Q\rangle = \langle P\,|\,u(Q)\rangle\ ,$$

donc u est un endomorphisme symétrique pour le produit scalaire considéré.

> ✓ Il ne servait à rien d'écrire la matrice de u dans la base canonique de $\mathbb{R}_n[X]$ pour montrer que u est symétrique, car cette base n'est pas orthonormale pour le produit scalaire considéré !

• u étant un endomorphisme symétrique, ses valeurs propres sont réelles. Si λ est une de ces valeurs propres et si P est un vecteur propre associé, on a $u(P) = \lambda P$ donc la relation :

$$\langle u(P)\,|\,P\rangle = -\int_{-1}^{1}(1-t^2)P'(t)^2\,\mathrm{d}t$$

implique $\lambda\,\|P\|^2 \leqslant 0$, d'où $\lambda \leqslant 0$ puisque $\|P\| \neq 0$.

> ✓ En écrivant la matrice de u dans la base canonique, qui est triangulaire supérieure, on trouve que ces valeurs propres sont les entiers $-k(k+1)$ pour $k \in [\![0\,;n]\!]$.

Exercice 6.4

Si ${}^tA = A^2$ alors $A^4 = {}^tA^2 = A$ donc le polynôme $X^4 - X$ est annulateur de A. Ce polynôme étant scindé à racines simples dans $\mathbb{C}[X]$, A est diagonalisable dans $\mathcal{M}_2(\mathbb{C})$, de valeurs propres incluses dans $\{0,1,\mathrm{j},\mathrm{j}^2\}$.

Deux cas se présentent.

• Si j est valeur propre de A, A étant réelle, $\mathrm{j}^2 = \bar{\mathrm{j}}$ est aussi valeur propre de A. A est donc semblable dans $\mathcal{M}_2(\mathbb{C})$ à $\mathrm{diag}(\mathrm{j},\mathrm{j}^2)$ donc $A^2 = A^{-1}$ donc ${}^tA = A^{-1}$: A est une matrice orthogonale. Puisque $\det A = \mathrm{j}.\mathrm{j}^2 = 1$, c'est une matrice de rotation ; et puisque $A^3 = \mathrm{I}_3$, l'angle de cette rotation est $\pm\dfrac{2\pi}{3}$.

• Sinon, les valeurs propres de A sont incluses dans $\{0,1\}$, et A est donc semblable dans $\mathcal{M}_2(\mathbb{C})$ à une matrice diagonale avec 0 et/ou 1 sur la diagonale. Il en résulte $A^2 = A$, donc A est la matrice d'une projection, et puisque $A = A^2 = {}^tA$, A est symétrique et il s'agit d'une projection orthogonale.

• Dans les deux cas, la réciproque est facile à établir.

Exercice 6.5

Notons $S = {}^tAA$. Puisque A et ${}^tA$ commutent, on a :

$$\forall\, k \in \mathbb{N},\ S^k = ({}^tA)^k A^k\ ,$$

et puisque A est nilpotente, on en déduit que S est nilpotente : il existe $p \in \mathbb{N}^*$ tel que $S^p = 0_n$.

Mais S est une matrice symétrique réelle (${}^tS = S$ se vérifie facilement) donc d'après le théorème spectral elle est semblable à une matrice de la forme $\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$.

S^p est donc semblable à $\mathrm{diag}(\lambda_1^p,\ldots,\lambda_n^p)$ et puisque $S^p = 0_n$ on en déduit $\lambda_i^p = 0$ pour tout i, donc $\lambda_i = 0$ pour tout i et finalement $S = 0$.

En conclusion, on a ${}^tAA = 0$. Mais dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, $\|A\|^2 = \mathrm{tr}({}^tAA)$. On en déduit $\|A\| = 0$ donc $A = 0_n$.

Exercice 6.6

1) Soient $x \in \mathrm{Ker}\,v$ et $y \in \mathrm{Im}\,v$; alors $v(x) = 0$ soit $u(x) = x$ et il existe $z \in E$ tel que $y = v(z) = z - u(z)$ donc on a :

$$\begin{aligned}\langle x\,|\,y\rangle &= \langle x\,|\,z-u(z)\rangle\\ &= \langle x\,|\,z\rangle - \langle u(x)\,|\,u(z)\rangle = 0\\ &\quad\ (\text{car } u \text{ isométrie})\,.\end{aligned}$$

On a donc $\mathrm{Ker}\,v \subset (\mathrm{Im}\,v)^\perp$, et le théorème du rang permet de conclure.

2) • Si $x \in \mathrm{Ker}\,v$ on a $u(x) = x$ et

$$\frac{1}{m}\sum_{k=0}^{m-1}u^k(x) = x = p(x)\,.$$

• Si $x \in \mathrm{Im}\,v$, il existe $z \in E$ tel que $x = v(z) = z - u(z)$ et on a :

$$\begin{aligned}\frac{1}{m}\sum_{k=0}^{m-1}u^k(x) &= \frac{1}{m}\sum_{k=0}^{m-1}u^k(z-u(z))\\ &= \frac{1}{m}\sum_{k=0}^{m-1}u^k(z) - u^{k+1}(z)\\ &= \frac{1}{m}\big(z-u^m(z)\big)\,.\end{aligned}$$

De plus,

$$\left\|\frac{1}{m}\big(z-u^m(z)\big)\right\| \leqslant \frac{1}{m}\left(\|z\|+\|u^m(z)\|\right)$$
$$\leqslant \frac{2}{m}\|z\| \quad (\text{car } u \text{ isométrie}).$$

On en déduit que :

$$\frac{1}{m}\sum_{k=0}^{m-1} u^k(x) \underset{m\to+\infty}{\longrightarrow} 0 = p(x).$$

• Ainsi l'affirmation est établie pour $x \in \operatorname{Ker} v$ et pour $x \in \operatorname{Im} v$; ces deux sous-espaces vectoriels étant supplémentaires d'après **1)**, cela établit le résultat pour tout $x \in E$ par linéarité.

Exercice 6.7

On munit $\mathbb{R}^n$ (que l'on confond avec $\mathcal{M}_{n,1}(\mathbb{R})$) de sa structure euclidienne canonique. On notera f l'endomorphisme de $\mathbb{R}^n$ dont M est la matrice dans la base canonique.
Soit $u = (1,1,\ldots,1)$. On a :

$$\langle u \,|\, f(u)\rangle = \left\langle \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} \middle| \begin{pmatrix}\sum_{j=1}^n m_{1j}\\ \vdots \\ \sum_{j=1}^n m_{nj}\end{pmatrix} \right\rangle$$
$$= \sum_{i=1}^n \sum_{j=1}^n m_{ij},$$

d'où par l'inégalité de Cauchy-Schwarz :

$$\left|\sum_{i=1}^n \sum_{j=1}^n m_{ij}\right| \leqslant \|u\| \cdot \|f(u)\|.$$

De plus, M est orthogonale, donc f est une isométrie d'où $\|f(u)\| = \|u\| = \sqrt{n}$, et on en déduit :

$$\left|\sum_{i=1}^n \sum_{j=1}^n m_{ij}\right| \leqslant n.$$

• Pour tout $j \in [\![1\,;n]\!]$, $\sum_{i=1}^n m_{ij}^2 = 1$ (norme du j^{e} vecteur colonne de M), d'où

$$\forall\,(i,j) \in [\![1\,;n]\!]^2,\ |m_{ij}| \leqslant 1,$$

ce qui implique : $m_{ij}^2 \leqslant |m_{ij}|$.
On en déduit :

$$\sum_{j=1}^n \sum_{i=1}^n m_{ij}^2 \leqslant \sum_{i=1}^n \sum_{j=1}^n |m_{ij}|,$$

et puisque $\sum_{j=1}^n \sum_{i=1}^n m_{ij}^2 = \sum_{j=1}^n 1 = n$, on obtient :

$$n \leqslant \sum_{i=1}^n \sum_{j=1}^n |m_{ij}|.$$

• Pour tout $(i,j) \in [\![1\,;n]\!]^2$, posons $|m_{ij}| = \varepsilon_{ij} m_{ij}$, avec $\varepsilon_{ij} \in \{-1,1\}$.

L'inégalité de Cauchy-Schwarz pour le produit scalaire canonique dans $\mathbb{R}^{n^2}$ donne alors :

$$\sum_{i,j} |m_{ij}| \leqslant \left(\sum_{i,j} \varepsilon_{ij}^2\right)^{\frac{1}{2}} \left(\sum_{i,j} m_{ij}^2\right)^{\frac{1}{2}}$$

et puisque $\sum_{i,j} \varepsilon_{ij}^2 = n^2$ (n^2 termes égaux à 1) et que $\sum_{i,j} m_{ij}^2 = n$ (matrice orthogonale), on en déduit la dernière inégalité cherchée.

Exercice 6.8

Dans cet exercice, on confondra comme d'habitude les matrices avec les endomorphismes qui leur sont canoniquement associés.

1) • Soit A une matrice symétrique positive, et soit $\lambda \in \operatorname{Sp} A$. On sait déjà que λ est réel d'après le théorème spectral.
De plus, si X est un vecteur propre associé, on a :

$${}^{\text{t}}XAX = {}^{\text{t}}X(\lambda X) = \lambda \|X\|^2,$$

et puisque $\|X\| \neq 0$, la propriété ${}^{\text{t}}XAX \geqslant 0$ implique $\lambda \geqslant 0$.

On a donc montré que $\operatorname{Sp} A \subset \mathbb{R}_+$.

• Réciproquement, soit A une matrice symétrique dont toutes les valeurs propres sont positives. On sait qu'il existe une base orthonormale de $\mathcal{M}_{n,1}(\mathbb{R})$ formée de vecteurs propres $(V_1,\ldots,V_n)$ de A. Soient $\lambda_1,\ldots,\lambda_n$ les valeurs propres associées, par hypothèse positives.
Tout vecteur X de $\mathcal{M}_{n,1}(\mathbb{R})$ s'écrit sous la forme :

$$X = \sum_{i=1}^n x_i V_i.$$

On a alors :

$$
\begin{aligned}
{}^tXAX &= \left(\sum_{i=1}^{n} x_i\,{}^tV_i\right) A \left(\sum_{j=1}^{n} x_j V_j\right) \\
&= \sum_{1\leqslant i,j\leqslant n} x_i x_j\,{}^tV_i A V_j \\
&= \sum_{1\leqslant i,j\leqslant n} x_i x_j \lambda_j\,{}^tV_i V_j\,.
\end{aligned}
$$

La base $(V_1,\dots,V_n)$ étant orthonormale, on a ${}^tV_iV_j = \langle V_i \,|\, V_j\rangle = \delta_{ij}$ donc il reste :

$$
{}^tXAX = \sum_{j=1}^{n} \lambda_j x_j^2\,,
$$

et puisque les λ_j sont positifs, il en est de même de ${}^tXAX$.

On a donc établi l'implication réciproque.

• Le cas des matrices symétrique définies positives se traite de la même façon, en remplaçant l'inégalité large $\geqslant 0$ par l'inégalité stricte > 0.

2) • Supposons qu'il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A = {}^tBB$.

Alors A est clairement symétrique et pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$
{}^tXAX = {}^tX\,{}^tBBX = \|BX\|^2 \geqslant 0\,,
$$

donc $A \in \mathcal{S}_n^+(\mathbb{R})$.

• Réciproquement, soit $A \in \mathcal{S}_n^+(\mathbb{R})$. A est orthodiagonalisable car symétrique réelle :

$$
A = PD\,{}^tP\,,
$$

avec $P \in \mathcal{O}_n(\mathbb{R})$ et $D = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$ diagonale. Puisque les valeurs propres de A sont positives d'après la 1re question, on peut écrire $D = \Delta^2$ avec $\Delta = \operatorname{diag}\left(\sqrt{\lambda_1},\dots,\sqrt{\lambda_n}\right)$. On aura alors :

$$
A = P(\Delta^2)\,{}^tP = (P\Delta)\left(\Delta\,{}^tP\right) = {}^tBB
$$

avec $B = \Delta\,{}^tP$ (puisque ${}^t\Delta = \Delta$).

• On démontre de la même façon que $A \in \mathcal{S}_n(\mathbb{R})$ est définie positive si et seulement si il existe une matrice $B \in \mathcal{M}_n(\mathbb{R})$, *inversible*, telle que $A = {}^tBB$.

3) • Soit $A \in \mathcal{S}_n^+(\mathbb{R})$. En reprenant les notations et calculs précédents, on peut aussi écrire :

$$
A = \left(P\Delta\,{}^tP\right)\left(P\Delta\,{}^tP\right)
$$

puisque ${}^tPP = \mathrm{I}_n$. La matrice $R = P\Delta\,{}^tP$ est alors symétrique et vérifie bien $R^2 = A$. De plus, ses valeurs propres sont les coefficients diagonaux de Δ, donc sont positives. Donc $R \in \mathcal{S}_n^+(\mathbb{R})$.

• Il reste à démontrer que R est unique.

Soit donc $S \in \mathcal{S}_n^+(\mathbb{R})$ telle que $S^2 = A$. Pour toute valeur propre λ de A (positive), on a :

$$
\operatorname{Ker}(S - \sqrt{\lambda}\mathrm{I}_n) \subset \operatorname{Ker}(A - \lambda \mathrm{I}_n) \quad (1)
$$

puisque

$$
SX = \sqrt{\lambda}X \implies S^2X = \lambda X\,.
$$

D'autre part, S commute avec A donc $\operatorname{Ker}(A - \lambda\mathrm{I}_n)$ est stable par S. Il en résulte que l'endomorphisme induit par S sur $\operatorname{Ker}(A - \lambda\mathrm{I}_n)$ est diagonalisable (puisque S est symétrique). Si μ en est une valeur propre et si $X \in \operatorname{Ker}(A - \lambda\mathrm{I}_n)$ est un vecteur propre associé, on a $SX = \mu X$ d'où

$$
\mu^2 X = S^2X = AX = \lambda X\,,
$$

donc $\lambda = \mu^2$ puis $\mu = \sqrt{\lambda}$ puisque les valeurs propres de S sont positives.
La seule valeur propre de l'endomorphisme induit par S sur $\operatorname{Ker}(A - \lambda\mathrm{I}_n)$ est donc $\sqrt{\lambda}$ et comme cet endomorphisme est diagonalisable, on a $\operatorname{Ker}(A-\lambda\mathrm{I}_n) \subset \operatorname{Ker}(S-\sqrt{\lambda}\mathrm{I}_n)$. Finalement :

$$
\forall\,\lambda \in \operatorname{Sp} A, \operatorname{Ker}(S-\sqrt{\lambda}\mathrm{I}_n) = \operatorname{Ker}(A-\lambda\mathrm{I}_n).
$$

L'endomorphisme canoniquement associé à S est donc nécessairement égal à $\sqrt{\lambda}\mathrm{Id}$ sur chaque sous-espace propre $E_\lambda(a)$ de l'endomorphisme canoniquement associé à A. Ces sous-espaces propres étant supplémentaires, cela le détermine de façon unique.

4) • Soit $S \in \mathcal{S}_n(\mathbb{R})$. Il existe donc $P \in \mathcal{O}_n(\mathbb{R})$ et D diagonale telles que $S = PDP^{-1}$. On a vu en **[SF4.29]** que l'on a alors :

$$
\exp(S) = P\exp(D)P^{-1}
$$

donc $\exp(S)$ est symétrique (car $P^{-1} = {}^tP$) et à valeurs propres strictement positives (ce sont les e^λ pour $\lambda \in \operatorname{Sp}(S)$), donc $\exp(S) \in \mathcal{S}_n^{++}(\mathbb{R})$.

• Si $U \in \mathcal{S}_n^{++}(\mathbb{R})$, alors $U = P\Delta P^{-1}$ avec P orthogonale et Δ diagonale à éléments diagonaux strictement positifs. Ces éléments diagonaux peuvent donc s'écrire e^{λ_i} avec $\lambda_i \in \mathbb{R}$, donc il existe D diagonale telle que $\Delta = \exp(D)$, et $U = \exp(S)$ avec $S = PDP^{-1}$.

S est symétrique réelle, ce qui montre que l'application exp est surjective de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.

• Soient maintenant S et S' symétriques réelles telles que $\exp(S) = \exp(S')$.

– Nous avons vu en **[SF4.30]** que l'exponentielle d'une matrice M est un polynôme en M. Donc $\exp(S')$ est un polynôme en S et par suite commute avec S.

– D'autre part, en diagonalisant S' et en considérant un polynôme Q tel que $Q(\mathrm{e}^{\lambda}) = \lambda$ pour toute valeur propre λ de S' (polynômes d'interpolation de Lagrange par exemple), on montre que $Q(\exp(S')) = S'$. Comme S commute avec $\exp(S')$, il en résulte que S commute avec S'.

– Puisque S et S' sont diagonalisables et commutent, elles possèdent une base commune de diagonalisation (voir **[SF4.7]** page 299), c'est-à-dire qu'il existe $P \in \mathrm{GL}_n(\mathbb{R})$ et D, D' diagonales telles que :

$$S = PDP^1 \quad \text{et} \quad S' = PD'P^{-1}\,.$$

La relation $\exp(S) = \exp(S')$ implique alors $\exp(D) = \exp(D')$ et puisque ces matrices sont réelles, il en résulte $D = D'$, puis $S = S'$.

La fonction exp est donc injective, et finalement bijective, de $\mathcal{S}_n(\mathbb{R})$ sur $\mathcal{S}_n^{++}(\mathbb{R})$.

Exercice 6.9

• Dans le cas trivial $H = H'$, on a $s = s'$ donc $s \circ s' = s' \circ s = \mathrm{Id}_E$.

• On supposera donc $H \neq H'$. Soient u et u' des vecteurs unitaires orthogonaux à H et H' respectivement.
On suppose que $s \circ s' = s' \circ s$.
On a $s(u) = -u$ donc

$$s'(u) = s'[s(-u)] = s[s'(-u)] = -s[s'(u)]$$

ce qui implique que $s'(u)$ est colinéaire à u. Puisque $s'(u)$ est unitaire (car s' isométrie) on a $s'(u) = \pm u$; le cas $s'(u) = -u$ étant impossible car cela impliquerait u et u' colinéaires, on a $s'(u) = u$ donc $u \in H'$: u est orthogonal à u'.
Ainsi $H^{\perp} \subset H'$ c'est-à-dire que H et H' sont perpendiculaires (ou encore $H^{\perp}$ et $H'^{\perp}$ sont orthogonaux).

• Réciproquement supposons H et H' perpendiculaires.

Alors $(H \cap H')^{\perp} = H^{\perp} \overset{\perp}{\oplus} H'^{\perp}$. On pourra donc trouver une base orthonormale formée de la réunion d'une base orthonormale de $H \cap H'$, d'un vecteur unitaire de $H^{\perp}$ et d'un vecteur unitaire de $H'^{\perp}$. Dans cette base, les matrices de s et s' sont diagonales (avec $n-1$ termes égaux à 1 et un terme égal à -1 sur la diagonale), donc commutent ce qui prouve que $s \circ s' = s' \circ s$.

• Conclusion : s et s' commutent si et seulement si $H = H'$ ou si H et H' sont perpendiculaires.

Exercice 6.10

1) On utilise ici une idée similaire à celle signalée en **[SF5.4]** page 337 (utilisation des identités de polarisation).
Si u vérifie la condition de l'énoncé, alors, pour tous x, $y \in E$:

$$\langle u(x+y)\,|\,x+y\rangle = 0\,,$$

d'où, en utilisant la linéarité de u et la bilinéarité du produit scalaire :

$$\underbrace{\langle u(x)\,|\,x\rangle}_{=0} + \langle u(x)\,|\,y\rangle + \langle u(y)\,|\,x\rangle + \underbrace{\langle u(y)\,|\,y\rangle}_{=0} = 0\,,$$

ce qui donne le résultat voulu.

2) Cette propriété se démontre exactement comme la propriété correspondante du cours pour les endomorphismes symétriques !

3) Ici aussi, on adapte la démonstration du cours où l'on a montré que le spectre d'une matrice symétrique réelle est inclus dans $\mathbb{R}$.
Soit donc $\lambda \in \mathbb{C}$ une valeur propre de A ; il existe $X \in \mathcal{M}_{n,1}(\mathbb{C})$, non nul, tel que $AX = \lambda X$. On a alors, d'une part :

$${}^{\mathrm{t}}\overline{X}AX = \lambda\,{}^{\mathrm{t}}\overline{X}X\,,$$

et d'autre part, puisque ${}^{\mathrm{t}}\overline{X}AX$ est un nombre complexe et que la matrice A est à coefficients réels :

$${}^{\mathrm{t}}\overline{X}AX = {}^{\mathrm{t}}\big({}^{\mathrm{t}}\overline{X}AX\big) = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}A\overline{X} = -{}^{\mathrm{t}}XA\overline{X}$$
$$= -{}^{\mathrm{t}}X(\overline{AX}) = -{}^{\mathrm{t}}X(\overline{\lambda X}) = -\overline{\lambda}\,{}^{\mathrm{t}}X\overline{X}$$

d'où $\lambda {}^t\overline{X}X = -\overline{\lambda}{}^t X\overline{X}$, et puisque ${}^t\overline{X}X = {}^tX\overline{X} = \sum_{i=1}^{n} |x_i|^2 \neq 0$, on en déduit $\lambda = -\overline{\lambda}$, c'est-à-dire λ imaginaire pur.

4) a) Soit F un sous-espace vectoriel de E stable par u, et soit $y \in F^\perp$. Pour tout $x \in F$, on a $u(x) \in F$ donc $\langle u(x) \,|\, y\rangle = 0$.
u étant antisymétrique, on en déduit $\langle x \,|\, u(y)\rangle = 0$, et cela étant valable pour tout $x \in F$, on a $u(y) \in F^\perp$.
$F^\perp$ est donc stable par u.

b) • Notons déjà que si A est orthogonalement semblable à une matrice D de la forme donnée, elle est antisymétrique puisqu'alors :

$$A = PD{}^tP \text{ avec } P \in \mathcal{O}_n(\mathbb{R})$$
$$\Longrightarrow {}^tA = P{}^tD{}^tP = P(-D){}^tP = -A\,.$$

Il s'agit donc de démontrer la réciproque, c'est à dire montrer que tout endomorphisme antisymétrique de E possède une matrice de la forme voulue dans une base orthonormale.

• Commençons par examiner le cas où u est un endomorphisme antisymétrique bijectif de E. Si $n = \dim E$, puisque la matrice A de u dans une base orthonormale est antisymétrique, on a :

$$\det A = \det {}^tA = \det(-A) = (-1)^n \det A\,,$$

d'où $(-1)^n = 1$ puisque $\det A \neq 0$: n est nécessairement pair.

Posons alors $n = 2m$, et établissons le résultat demandé par récurrence sur m.

– Pour $m = 1$, la matrice de u dans une base orthonormale (e_1,e_2) est antisymétrique réelle inversible donc de la forme $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ où $a \in \mathbb{R}^*$. Et quitte à changer (e_1,e_2) en $(e_1, -e_2)$, on peut supposer $a > 0$.
Le résultat est donc démontré dans ce cas.

– Supposons le résultat démontré pour tout endomorphisme antisymétrique inversible d'un espace vectoriel de dimension $2(m-1)$, et soit u un automorphisme antisymétrique d'un espace vectoriel E de dimension $2m$.
Il est facile de vérifier que u^2 est alors un endomorphisme symétrique de E, bijectif. u^2 est donc diagonalisable ; soit $\lambda \in \mathbb{R}$ une valeur propre non nulle de u^2, et e_1 un vecteur propre associé, que l'on supposera unitaire.
Puisque $\langle u(e_1) \,|\, e_1\rangle = 0$, $u(e_1)$ est orthogonal à e_1, et est non nul puisque u est bijective. Posons alors $e_2 = \dfrac{u(e_1)}{\|u(e_1)\|}$, de sorte que la famille (e_1,e_2) est orthonormale.
On a alors $u(e_1) = ae_2$ avec $a = \|u(e_1)\| > 0$, puis

$$u(e_2) = \frac{1}{a}u^2(e_1) = \frac{\lambda}{a}e_1 = be_1\,.$$

Mais alors la relation :

$$0 = \langle u(e_1+e_2) \,|\, e_1+e_2\rangle$$
$$= \langle ae_2 + be_1 \,|\, e_1+e_2\rangle = a+b$$

implique $b = -a$, soit $u(e_2) = -ae_1$.
Le plan $P = \mathrm{Vect}(e_1,e_2)$ est donc stable par u, et la matrice de l'endomorphisme induit par u dans la base orthonormale (e_1,e_2) de ce plan est $\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$ avec $a > 0$.
D'après la question **4)a)**, $P^\perp$ est un espace vectoriel de dimension $2(m-1)$ stable par u ; l'endomorphisme induit par u sur $P^\perp$ est antisymétrique bijectif, il existe donc d'après l'hypothèse de récurrence une base orthonormale de ce sous-espace où sa matrice a la forme voulue.
La réunion de (e_1,e_2) et de cette base fournit donc une base orthonormale de E où la matrice de u est de la forme voulue, ce qui achève la récurrence.

• Dans le cas général, soit u un endomorphisme antisymétrique de E.
Puisque $E = \operatorname{Ker} u \oplus (\operatorname{Ker} u)^\perp$, la restriction de u à $(\operatorname{Ker} u)^\perp$ est un isomorphisme de $(\operatorname{Ker} u)^\perp$ sur $\operatorname{Im} u$, d'après le théorème d'isomorphisme.
Puisque $(\operatorname{Ker} u)^\perp$ est stable par u, on a donc

$$\operatorname{Im} u = u\big((\operatorname{Ker} u)^\perp\big) \subset (\operatorname{Ker} u)^\perp\,,$$

donc $\operatorname{Im} u = (\operatorname{Ker} u)^\perp$ par égalité des dimensions, et l'endomorphisme induit par u sur $(\operatorname{Ker} u)^\perp$ est un automorphisme antisymétrique de ce sous-espace.

On peut donc lui appliquer le résultat précédent, et en faisant la réunion de la base orthonormale obtenue avec une base

orthonormale de $\operatorname{Ker} u$, on obtient le résultat.

Exercice 6.11

1) Par hypothèse, il existe une base $\mathscr{B}$ telle que $A = \mathcal{M}_{\mathscr{B}}(u)$ est normale. Si $\mathscr{B}'$ est une autre base orthonormale de E, si P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ et si $A' = \mathcal{M}_{\mathscr{B}'}(u)$, on a alors $A' = {}^tPAP$ avec ${}^tPP = \mathrm{I}_n$ d'où :

$$\begin{aligned} {}^tA'A' &= ({}^tP{}^tAP)({}^tPAP) \\ &= {}^tP({}^tAA)P = {}^tP(A{}^tA)P = A'{}^tA' , \end{aligned}$$

donc A' est aussi normale.

2) On suppose que F est un sous-espace vectoriel de E stable par u. On peut alors construire une base orthonormale formée de la réunion d'une base orthonormale de F et d'une base orthonormale de $F^\perp$. On sait que la matrice de u dans cette base s'écrit par blocs : $A = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$.
Puisque A est normale, on a :

$$\begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \times \begin{bmatrix} {}^tX & 0 \\ {}^tY & {}^tZ \end{bmatrix} = \begin{bmatrix} {}^tX & 0 \\ {}^tY & {}^tZ \end{bmatrix} \times \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix},$$

ce qui implique :

$$X{}^tX + Y{}^tY = {}^tXX ,$$

d'où $\operatorname{tr}(X{}^tX + Y{}^tY) = \operatorname{tr}({}^tXX)$, et enfin $\operatorname{tr}(Y{}^tY) = 0$ puisque $\operatorname{tr}({}^tXX) = \operatorname{tr}(X{}^tX)$. Cela implique, d'après un résultat classique redémontré en **[SF3.1]** page 249, que la matrice Y est nulle. Le sous-espace vectoriel $F^\perp$ est donc aussi stable par u.

3) Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ une matrice normale. La relation $A{}^tA = {}^tAA$ équivaut à :

$$\begin{cases} b^2 - c^2 = 0 \\ ac + bd = ab + cd \end{cases}$$

soit à :

$$\begin{cases} (b-c)(b+c) = 0 \\ (b-c)(a-d) = 0 \end{cases}$$

Si $b = c$, A est symétrique réelle donc orthodiagonalisable , c'est-à-dire orthogonalement semblable à une matrice de la forme $\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.
Sinon, $a = d$ et $b = -c$ donc $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

4) • Si A est orthogonalement semblable à une matrice de la forme voulue, il n'est pas bien difficile de vérifier qu'elle commute avec sa transposée.

• Pour la réciproque, il faut donc établir que tout endomorphisme normal possède dans une base orthonormale une matrice de la forme voulue.
On démontre ce résultat par récurrence forte sur $n = \dim E$.

- Pour $n = 1$, le résultat est immédiat, et il vient d'être établi pour $n = 2$.

- Supposons le résultat établi pour tout endomorphisme normal d'un espace euclidien de dimension $\leqslant n-1$, et soit u un endomorphisme normal d'un espace E de dimension n.
  D'après le résultat mentionné dans l'énoncé, il existe un sous-espace vectoriel F de dimension 1 ou 2 stable par u.

  - Si $\dim F = 1$ alors l'hyperplan $F^\perp$ est stable par u, et l'endomorphisme induit par u sur $F^\perp$ est encore normal (pour le vérifier, considérer une base orthonormale adaptée à la décomposition $E = F \oplus F^\perp$, et écrire la matrice de u dans cette base).
    Il existe donc une base orthonormale de $F^\perp$ où la matrice de u a la forme voulue, et il suffit d'ajouter à cette base un vecteur unitaire de F pour conclure.

  - Si $\dim F = 2$, on considère les endomorphismes induits par u sur F et sur $F^\perp$; ce sont encore des endomorphismes normaux, et on applique alors le résultat à l'ordre 2 et à l'ordre $n-2$ pour construire une base orthonormale qui convient.

- Cela achève la récurrence, et établit le résultat demandé.

Exercice 7.1

• *N norme*

La définition de N a bien un sens (l'intégrale existe puisqu'on suppose f' continue).

Pour montrer que N est une norme, on considère l'application :

$$\begin{aligned}\varphi\colon E\times E &\longrightarrow \mathbb{R}\\ (f,g) &\longmapsto f(0)g(0)+\int_0^1 f'(t)g'(t)\,\mathrm{d}t\end{aligned}$$

φ est une forme bilinéaire symétrique sur E (facile), positive ($\forall f\in E,\ \varphi(f,f)\geqslant 0$) et définie car :

$$\begin{aligned}\varphi(f,f)=0 &\\ &\Longrightarrow f(0)=0 \text{ et } \int_0^1 f'^2(t)\,\mathrm{d}t=0\\ &\Longrightarrow f(0)=0 \text{ et } f'=0 \Longrightarrow f=0\,.\end{aligned}$$

C'est donc un produit scalaire sur E, et $N : f\mapsto\sqrt{\varphi(f,f)}$ est la norme euclidienne associée.

• *N' norme*

La définition de N' a bien un sens puisque, f et f' étant continues sur $[0\,;1]$ y sont bornées. La démonstration du fait que N' est une norme est aisée en utilisant directement le fait que $\|\ \|_\infty$ est une norme d'après le cours.

• *Comparaison de N et N'*

Pour $f\in E$, on a :

$$\begin{aligned}N(f)^2&=f(0)^2+\int_0^1 f'^2(t)\,\mathrm{d}t\\ &\leqslant \|f\|_\infty^2+\|f'\|_\infty^2\end{aligned}$$

$$\begin{aligned}\text{donc : } N(f)&\leqslant\sqrt{\|f\|_\infty^2+\|f'\|_\infty^2}\\ &\leqslant \|f\|_\infty+\|f'\|_\infty=N'(f)\,.\end{aligned}$$

Cependant, N et N' ne sont pas équivalentes ; en effet, considérons, pour $t\in[0\,;1]$ et $n\in\mathbb{N}^*$, $f_n(t)=t^n$.
Alors $f_n\in E$, $N(f_n)=\dfrac{n}{\sqrt{2n-1}}$
et $N'(f_n)=1+n$ donc $\lim\limits_{n\to\infty}\dfrac{N'(f_n)}{N(f_n)}=+\infty$
et le rapport $\dfrac{N'}{N}$ n'est pas borné sur $E\backslash\{0\}$.

• *Comparaison de N et de $\|\ \|_\infty$*

En considérant le même exemple que ci-dessus, on peut voir que le rapport $\dfrac{N}{\|\ \|_\infty}$ n'est pas borné.

Cependant, on remarque que

$$\forall t\in[0\,;1],\ f(t)=f(0)+\int_0^t f'(u)\,\mathrm{d}u\,.$$

Donc :

$$\begin{aligned}|f(t)|&\leqslant|f(0)|+\int_0^t\big|f'(u)\big|\,\mathrm{d}u\\ &\leqslant|f(0)|+\int_0^1\big|f'(u)\big|\,\mathrm{d}u\\ &\leqslant|f(0)|+\sqrt{\int_0^1 1\,\mathrm{d}u\int_0^1 f'^2(u)\,\mathrm{d}u}\end{aligned}$$

d'après l'inégalité de Cauchy-Schwarz, et, puisque $a+b\leqslant\sqrt{2}\sqrt{a^2+b^2}$ pour tous $a,\,b\in\mathbb{R}$, on en déduit :

$$\forall t\in[0\,;1],\ |f(t)|\leqslant\sqrt{2}N(f)\,,$$

d'où

$$\|f\|_\infty\leqslant\sqrt{2}N(f)\,.$$

• *Comparaison de N' et de $\|\ \|_\infty$*

L'inégalité $\|f\|_\infty\leqslant N'(f)$ pour toute f est immédiate.
Et le même exemple, encore, permet de montrer facilement que N' et $\|\ \|_\infty$ ne sont pas équivalentes.

Exercice 7.2

• Montrons d'abord que N est une norme sur $\mathcal{M}_n(\mathbb{K})$.

– L'application N est bien définie sur $\mathcal{M}_n(\mathbb{K})$ et à valeurs réelles positives.

– Si $N(A)=0$ alors pour tout $i\in[\![1\,;n]\!]$, $\sum\limits_{j=1}^n|a_{ij}|=0$ donc $a_{ij}=0$ pour tous i,j et $A=0_n$.

– Soit $\lambda\in\mathbb{K}$. L'égalité $N(\lambda A)=|\lambda|\,N(A)$ est immédiate puisque $|\lambda a_{ij}|=|\lambda|\,|a_{ij}|$ pour tous i, j.

– Soient enfin $A=(a_{ij})$ et $B=(b_{ij})$ deux matrices carrées, et $C=A+B$, $C=(c_{ij})$.

Pour tout $(i,j)\in[\![1\,;n]\!]^2$ on a $|c_{ij}|\leqslant|a_{ij}|+|b_{ij}|$ d'après l'inégalité triangulaire dans $\mathbb{K}$ d'où pour tout $i\in[\![1\,;n]\!]$:

$$\begin{aligned}\sum_{j=1}^n|c_{ij}|&\leqslant\sum_{j=1}^n|a_{ij}|+\sum_{j=1}^n|b_{ij}|\\ &\leqslant N(A)+N(B)\,,\end{aligned}$$

puis en prenant le max :

$$N(C)\leqslant N(A)+N(B)\,.$$

C'est l'inégalité triangulaire pour N, et cela achève de démontrer que N est une norme.

• Avec les mêmes notations que ci-dessus, soit $C = AB$. On a, pour tout $i \in [\![1\,;n]\!]$:

$$\begin{aligned}\sum_{j=1}^{n}|c_{ij}| &= \sum_{j=1}^{n}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right| \\ &\leqslant \sum_{j=1}^{n}\sum_{k=1}^{n}|a_{ik}b_{kj}| \\ &\leqslant \sum_{k=1}^{n}|a_{ik}|\underbrace{\left(\sum_{j=1}^{n}|b_{kj}|\right)}_{\leqslant N(B)} \\ &\leqslant N(B)\sum_{k=1}^{n}|a_{ik}| \leqslant N(B)N(A)\,.\end{aligned}$$

Ainsi : $N(AB) \leqslant N(A)N(B)$, donc N est une norme d'algèbre.

• Enfin, c'est une norme d'algèbre unitaire puisque $N(\mathrm{I}_n) = 1$.

Exercice 7.3

Nous allons démontrer que le complémentaire de U est un ouvert.

Soit donc $x \notin U$, et soit $A > \|x\|$.

Par définition de la limite infinie :

$$\exists\, n_0 \in \mathbb{N},\ n > n_0 \implies \|u_n\| > A + 1\,.$$

Ainsi, x et seulement un nombre fini de termes de la suite parmi u_0, …,u_{n_0} appartiennent à la boule fermée de centre 0 et de rayon $A+1$, et tous les autres termes de la suite sont à l'extérieur de cette boule.

Si l'on choisit maintenant $r > 0$ tel que :

$$r < \min\{\|u_0 - x\|, \ldots, \|u_{n_0} - x\|, 1\}\,,$$

ce qui est possible car x n'est pas l'un des u_i, la boule ouverte de centre x et de rayon r ne peut contenir aucun des termes de la suite : en effet, elle est incluse dans $\mathcal{B}_f(0,A+1)$ (car $\|x\| < A$ et $r < 1$, et d'après l'inégalité triangulaire) donc ne contient aucun des u_i pour $i > n_0$, et elle ne contient pas u_0, …, u_{n_0} par définition de r.

On a ainsi prouvé :

$$\forall\, x \in \complement_E^U,\ \exists\, r > 0 \text{ tel que } \mathcal{B}(x,r) \subset \complement_E^U\,,$$

ce qui prouve que $\complement_E^U$ est un ouvert et que U est fermé.

✓ On peut aussi démontrer ce résultat en utilisant la caractérisation séquentielle des fermés. Nous vous laissons y réfléchir.

Exercice 7.4

L'application $M \mapsto {}^{\mathrm{t}}M$ est linéaire sur $\mathcal{M}_p[\mathbb{R})$ qui est de dimension finie, donc elle est continue. Il en résulte que :

$$\lim_{n\to+\infty} A^n = B \implies \lim_{n\to+\infty} {}^{\mathrm{t}}A^n = {}^{\mathrm{t}}B\,.$$

Puisque A est antisymétrique, on a ${}^{\mathrm{t}}A = -A$ donc pour tout n :

$${}^{\mathrm{t}}A^{2n} = \left({}^{\mathrm{t}}A\right)^{2n} = A^{2n} \qquad (1)$$

et

$${}^{\mathrm{t}}A^{2n+1} = \left({}^{\mathrm{t}}A\right)^{2n+1} = -A^{2n+1} \qquad (2)$$

Les suites $\left({}^{\mathrm{t}}A^{2n}\right)_{n\in\mathbb{N}}$ et $\left({}^{\mathrm{t}}A^{2n+1}\right)_{n\in\mathbb{N}}$ étant des suites extraites de la suite $\left({}^{\mathrm{t}}A^{n}\right)_{n\in\mathbb{N}}$ convergent toutes deux vers ${}^{\mathrm{t}}B$, et de même les suites $\left(A^{2n}\right)_{n\in\mathbb{N}}$ et $\left(A^{2n+1}\right)_{n\in\mathbb{N}}$ convergent toutes deux vers B.
Des relations (1) et (2) on déduit alors par passage à la limite :

$$B = {}^{\mathrm{t}}B = -B\,,$$

d'où $B = 0_p$.

Exercice 7.5

En utilisant la formule du binôme :

$$\begin{aligned}&\exp(A) - \left(\mathrm{I}_n + \frac{1}{n}A\right)^n \\ &\quad= \sum_{k=0}^{+\infty}\frac{A^k}{k!} - \sum_{k=0}^{n}\binom{n}{k}\frac{A^k}{n^k} \\ &\quad= \sum_{k=0}^{n}\left(\frac{1}{k!} - \frac{\binom{n}{k}}{n^k}\right)A^k + \sum_{k=n+1}^{+\infty}\frac{A^k}{k!}\cdot\end{aligned}$$

Munissons $\mathcal{M}_p(\mathbb{K})$ d'une norme d'algèbre unitaire $\|\cdot\|$ (voir un exemple à l'exercice **7.2**). On a alors, par récurrence, $\left\|A^k\right\| \leqslant \|A\|^k$ pour tout entier k donc, en utilisant l'inégalité triangulaire :

$$\begin{aligned}&\left\|\exp(A) - \left(\mathrm{I}_n + \frac{1}{n}A\right)^n\right\| \\ &\quad\leqslant \sum_{k=0}^{n}\left|\frac{1}{k!} - \frac{\binom{n}{k}}{n^k}\right|\|A\|^k + \sum_{k=n+1}^{+\infty}\frac{\|A\|^k}{k!}\cdot\end{aligned}$$

Or pour tout entier k, le terme $\dfrac{1}{k!} - \dfrac{\binom{n}{k}}{n^k}$ est positif puisque $k!\binom{n}{k} = n(n-1)\cdots(n-k+1) \leqslant n^k$, donc :

$$\left|\frac{1}{k!} - \frac{\binom{n}{k}}{n^k}\right| = \frac{1}{k!} - \frac{\binom{n}{k}}{n^k}\,.$$

D'autre part, un calcul similaire au précédent donne :

$$\sum_{k=0}^{n}\left(\frac{1}{k!} - \frac{\binom{n}{k}}{n^k}\right)\|A\|^k + \sum_{k=n+1}^{+\infty}\frac{\|A\|^k}{k!} = \mathrm{e}^{\|A\|} - \left(1+\frac{1}{n}\|A\|\right)^n,$$

donc on a la majoration :

$$\left\|\exp(A) - \left(\mathrm{I}_n + \frac{1}{n}A\right)^n\right\| \leqslant \mathrm{e}^{\|A\|} - \left(1+\frac{1}{n}\|A\|\right)^n.$$

On sait que pour x réel positif :

$$\lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n = \mathrm{e}^x\,,$$

donc :

$$\lim_{n\to+\infty}\left(\mathrm{e}^{\|A\|} - \left(1+\frac{1}{n}\|A\|\right)^n\right) = 0\,,$$

et finalement :

$$\lim_{n\to+\infty}\left(\exp(A) - \left(\mathrm{I}_n + \frac{1}{n}A\right)^n\right) = 0\,.$$

Exercice 7.6

• $(i) \Longrightarrow (ii)$.
Notons λ_i $(1 \leqslant i \leqslant r)$ les valeurs propres distinctes de M dans $\mathbb{C}$. On suppose donc ici $|\lambda_i| < 1$ pour tout i.
On sait (voir par exemple **[SF4.19]** page 314) que M est semblable à une matrice de la forme :

$$T = \begin{bmatrix} \boxed{T_1} & & 0 \\ & \boxed{T_2} & \\ & \ddots & \\ 0 & & \boxed{T_r} \end{bmatrix}$$

où chaque T_i est triangulaire supérieure, de la forme $T_i = \lambda_i \mathrm{I}_{m_i} + N_i$ avec N_i nilpotente. On a donc $M = PTP^{-1}$ avec $P \in \mathrm{GL}_p(\mathbb{C})$, d'où pour tout n, $M^n = PT^nP^{-1}$. L'application $A \mapsto PAP^{-1}$ étant continue (car linéaire en dimension finie), montrer que $\lim\limits_{n\to+\infty} M^n = 0$ équivaut à montrer que $\lim\limits_{n\to+\infty} T^n = 0$.
D'autre part,

$$T^n = \begin{bmatrix} \boxed{T_1^n} & & 0 \\ & \boxed{T_2^n} & \\ & \ddots & \\ 0 & & \boxed{T_r^n} \end{bmatrix},$$

donc on est ramené à démontrer que $\lim\limits_{n\to+\infty} T_i^n = 0$ pour tout i.
Cela découle alors de la formule du binôme puisque, si N_i est nilpotente d'indice m, on a, pour $n \geqslant m$:

$$T_i^n = \sum_{k=0}^{m}\binom{n}{k}\lambda_i^{n-k}N^k\,,$$

et dans cette somme de $m+1$ termes, chaque terme $\binom{n}{k}\lambda_i^{n-k}$ tend vers 0 quand $n \to +\infty$ (en effet, $|\lambda_i| < 1$, $\binom{n}{k} \underset{n\to+\infty}{\sim} \dfrac{n^k}{k!}$ et par les croissances comparées).

• $(ii) \Longrightarrow (i)$.
Supposons $\lim\limits_{n\to+\infty} M^n = 0$, et soit $\lambda \in \mathbb{C}$ une valeur propre de M. Il existe donc $X \in \mathcal{M}_{p,1}(\mathbb{C})$, $X \neq 0$, tel que $MX = \lambda X$.
L'application $A \mapsto AX$ de $\mathcal{M}_p(\mathbb{C})$ dans $\mathcal{M}_{p,1}(\mathbb{C})$ étant continue (car linéaire en dimension finie) on a :

$$\lim_{n\to+\infty}(M^nX) = (\lim_{n\to+\infty} M^n)X = 0\,.$$

Mais pour tout $n \in \mathbb{N}$ on a $M^nX = \lambda^nX$; puisque $X \neq 0$ on en déduit $\lim\limits_{n\to+\infty}\lambda^n = 0$ donc $|\lambda| < 1$.

• $(ii) \Longrightarrow (iii)$.
Supposons $\lim\limits_{n\to+\infty} M^n = 0$. On vient de voir que toute valeur propre de M est de module strictement inférieur à 1 donc en particulier 1 n'est pas valeur propre de M, et $\mathrm{I}_p - M$ est inversible.
Pour tout entier n on a :

$$\left(\sum_{k=0}^{n}M^k\right)(\mathrm{I}_p - M) = \mathrm{I}_p - M^{n+1}\,,$$

d'où :

$$\left(\sum_{k=0}^{n}M^k\right) = (\mathrm{I}_p - M)^{-1}\left(\mathrm{I}_p - M^{n+1}\right)\,.$$

Par continuité on en déduit :

$$\lim_{n\to+\infty}\left(\sum_{k=0}^{n} M^k\right) = (\mathrm{I}_p - M)^{-1}\,.$$

La série $\sum M^k$ est donc convergente, de somme $(\mathrm{I}_p - M)^{-1}$.

• Pour terminer, l'implication $(iii) \implies (ii)$ est un résultat du cours sur les séries.

Exercice 7.7

• Soient x, $y \in \overline{C}$, et $\lambda \in [0\,;1]$.
Alors, par la définition séquentielle de l'adhérence, il existe des suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de C telles que :

$$x = \lim_{n\to+\infty} x_n \quad \text{et} \quad y = \lim_{n\to+\infty} y_n\,.$$

Par convexité de C, les éléments $\lambda x_n + (1-\lambda)y_n$ appartiennent à C, et, d'après les théorèmes usuels,

$$\lim_{n\to+\infty}(\lambda x_n + (1-\lambda)y_n) = \lambda x + (1-\lambda)y\,,$$

donc $\lambda x + (1-\lambda)y \in \overline{C}$, toujours à l'aide de la caractérisation séquentielle de l'adhérence.

Donc $\overline{C}$ est convexe.

• Soient x, $y \in \overset{\circ}{C}$, et $\lambda \in [0\,;1]$.
On pose $z = \lambda x + (1-\lambda)y$. Montrons que $z \in \overset{\circ}{C}$.

Par définition de l'intérieur, il existe une boule ouverte $\mathcal{B}(x,r') \subset C$ et une boule ouverte $\mathcal{B}(y,r'') \subset C$. Soit $r = \min(r',r'')$.

Nous vous laissons le soin de vérifier que :

$$\mathcal{B}(z,r) = \lambda\mathcal{B}(x,r) + (1-\lambda)\mathcal{B}(y,r)$$

(faites une figure), donc $\mathcal{B}(z,r) \subset C$ car C convexe.

Ainsi, $z \in \overset{\circ}{C}$ et $\overset{\circ}{C}$ est convexe.

Exercice 7.8

• u étant une suite bornée d'un espace vectoriel normé de dimension finie, le théorème de Bolzano-Weierstrass assure que l'ensemble A de ses valeurs d'adhérence n'est pas vide.

• Montrons que A est une partie bornée. Si $\ell \in A$, il existe une suite extraite $(u_{\varphi(n)})$ de limite ℓ. Par définition de la limite, on a en particulier :

$$\exists\, N \in \mathbb{N},\ n \geqslant N \implies \left\|u_{\varphi(n)} - \ell\right\| \leqslant 1\,.$$

Donc à l'aide de l'inégalité triangulaire :

$$\|\ell\| \leqslant \left\|u_{\varphi(N)} - \ell\right\| + \left\|u_{\varphi(N)}\right\| \leqslant \|u\|_\infty + 1\,,$$

donc A est inclus dans la boule $\mathcal{B}_f\left(0, \|u\|_\infty + 1\right)$, et est bornée.

• Montrons que A est un fermé, en montrant que son complémentaire est ouvert.
Soit $x \in \complement_E^A$. Puisque x n'est pas valeur d'adhérence de la suite u, il existe $r > 0$ tel que la boule ouverte $\mathcal{B}(x,r)$ ne contienne qu'un nombre fini de termes de la suite.
Pour tout $y \in \mathcal{B}(x,r)$, la boule ouverte $\mathcal{B}(y,r - \|x-y\|)$, incluse dans la précédente, ne contient donc elle aussi qu'un nombre fini de termes de la suite u, ce qui montre que y n'est pas valeur d'adhérence de cette suite.
Par conséquent, $\mathcal{B}(x,r) \subset \complement_E^A$, donc $\complement_E^A$ est un ouvert et A est fermée.

• En conclusion, A est une partie fermée bornée d'un espace vectoriel de dimension finie : c'est un compact.

Exercice 7.9

• Pour tout $(i,j) \in [\![1\,;n]\!]^2$, les applications $\varphi_{ij} : M \mapsto m_{ij}$ qui à une matrice M associent le terme d'indice (i,j) sont continues, car ce sont des applications linéaires sur un espace vectoriel normé de dimension finie.
Pour tout $i \in [\![1\,;n]\!]$, les applications $s_i : M \mapsto \sum_{j=1}^{n} m_{ij}$ sont également continues, comme somme d'applications continues.
On peut alors caractériser l'ensemble des matrices stochastiques comme étant l'ensemble formé des intersections :

$$\left(\bigcap_{(i,j)\in[\![1;n]\!]^2} \varphi_{ij}^{-1}\left([0\,;1]\right)\right) \bigcap \left(\bigcap_{i=1}^{n} s_i^{-1}(\{1\})\right).$$

Chacun des ensembles $\varphi_{ij}^{-1}\left([0\,;1]\right)$ et $s_i^{-1}(\{1\})$ étant un fermé comme image réciproque d'un fermé de $\mathbb{R}$ par une application

continue, l'ensemble des matrices stochastiques est l'intersection d'une famille de fermés, donc c'est un fermé.

✓ On pouvait ici aussi utiliser la caractérisation séquentielle d'un fermé en montrant que la limite d'une suite convergente de matrices stochastiques est encore une matrice stochastique.

• Si A est stochastique, on a $a_{ij} \in [0\,;1]$ pour tous i,j, donc $\|A\|_\infty \leqslant 1$. L'ensemble des matrices stochastiques est donc borné ; puisque c'est une partie fermée bornée d'un espace vectoriel de dimension finie, c'est un compact.

Exercice 7.10

1) Notons $U = \mathbb{C} \setminus \{a_1, \dots, a_p\}$, et soient a, $b \in U$. On peut construire construire un chemin de a à b inclus dans U de la façon suivante.
Soit D_a une droite du plan complexe passant par a et ne contenant aucun des a_i, et D_b une droite passant par b ne contenant aucun des a_i et non parallèle à D_a ; ces choix sont possibles puisqu'il n'existe qu'un nombre fini de droites passant par un point donné et l'un des a_i.
Soit $c = D_a \cap D_b$. Alors le chemin formé de la juxtaposition des segments $[a\,;c]$ et $[c\,;b]$ est un chemin continu reliant a et b et inclus dans U.
Cela prouve que U est connexe par arcs.

2) Soient A, $B \in \mathrm{GL}_n(\mathbb{C})$. Pour $z \in \mathbb{C}$ soit :

$$P(z) = \det(zA + (1-z)B)\,,$$

et soient $a_1, \dots, a_p$ les racines distinctes dans $\mathbb{C}$ de ce polynôme. Puisque $P(0) \neq 0$ et $P(1) \neq 0$ (car A et B sont inversibles), 0 et 1 appartiennent à $\mathbb{C} \setminus \{a_1, \dots, a_p\}$.
D'après ce qui précède, il existe un chemin continu $\varphi\colon t \in [0\,;1] \to \mathbb{C}$ tel que :

$$\varphi(0) = 0,\ \varphi(1) = 1,$$
$$\forall\, t \in [0\,;1],\ \varphi(t) \in \mathbb{C} \setminus \{a_1, \dots, a_p\}\,.$$

Alors l'application :

$$\gamma\colon t \in [0\,;1] \mapsto P\big(\varphi(t)\big)$$

est un chemin continu reliant A à B et qui est inclus dans $\mathrm{GL}_n(\mathbb{C})$.

Cela prouve que $\mathrm{GL}_n(\mathbb{C})$ est connexe par arcs.

Exercice 7.11

• Supposons f uniformément continue, et soit $\varepsilon > 0$. Par définition, il existe $\alpha > 0$ tel que, pour tous x, $y \in E$:

$$\|x - y\| < \alpha \implies \|f(x) - f(y)\| < \varepsilon.$$

Puisque $\lim\limits_{n\to+\infty} \|x_n - y_n\| = 0$, il existe $N \in \mathbb{N}$ tel que :

$$n \geqslant N \implies \|x_n - y_n\| < \alpha\,.$$

On en déduit que :

$$n \geqslant N \implies \|f(x_n) - f(y_n)\| < \varepsilon\,,$$

ce qui signifie que

$$\lim_{n\to+\infty} \|f(x_n) - f(y_n)\| = 0\,.$$

• Pour montrer la réciproque, raisonnons par l'absurde.
Supposons la propriété de l'énoncé vérifiée, et f non uniformément continue. Cela signifie qu'il existe $\varepsilon > 0$ tel que :

$$\forall\, \alpha > 0,\ \exists\, (x,y) \in E^2,$$
$$\|x - y\| < \alpha \text{ et } \|f(x) - f(y)\| \geqslant \varepsilon.$$

En particulier, pour tout $n \in \mathbb{N}^*$ il existe x_n et y_n dans E tels que :

$$\|x_n - y_n\| < \frac{1}{n} \text{ et } \|f(x_n) - f(y_n)\| \geqslant \varepsilon.$$

On a donc trouvé des suites $(x_n)_{n\in\mathbb{N}^*}$ et $(y_n)_{n\in\mathbb{N}^*}$ dans E telles que $\lim\limits_{n\to+\infty} \|x_n - y_n\| = 0$ et $\lim\limits_{n\to+\infty} \|f(x_n) - f(y_n)\| \neq 0$, ce qui est contradictoire.

Exercice 7.12

1) • Le résultat est immédiat si $r = n$ puisqu'alors l'ensemble considéré est égal à $\mathcal{M}_n(\mathbb{K})$ entier.

• Supposons donc $r < n$ pour la suite.
On sait qu'une matrice A est de rang strictement supérieur à r si et seulement si il existe une matrice carrée extraite de A, inversible et de taille strictement supérieure à r.
Par contraposition, on obtient que : $\operatorname{rg} A \leqslant r$ si et seulement si toute matrice carrée extraite de A et de taille strictement supérieure à r n'est pas inversible.

On définit alors une application f de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ par :

$$f(A) = \sum_{\substack{M \text{ extraite de } A \\ \text{taille}(M) > r}} |\det M| \,.$$

f est continue car c'est une somme de fonctions polynomiales des coefficients (composées avec la valeur absolue ou le module). D'après ce qui précède :

$$\operatorname{rg} A \leqslant r \iff f(A) = 0\,.$$

L'ensemble des matrices de rang inférieur ou égal à r est donc l'image réciproque par f continue du singleton $\{0\}$, qui est un fermé : c'est un fermé.

2) Notons :

$$E_r = \{A \in \mathcal{M}_n(\mathbb{K}) \mid \operatorname{rg} A = r\}$$
$$F_r = \{A \in \mathcal{M}_n(\mathbb{K}) \mid \operatorname{rg} A \leqslant r\}\,.$$

• On a $E_r \subset F_r$, et F_r est un fermé d'après ce qui précède, donc $\overline{E_r} \subset F_r$.

• Démontrons l'inclusion réciproque.
Si $A \in F_r$, on sait qu'il existe des matrices inversibles P et Q telles que $A = PJ_sQ$ où $J_s = \begin{bmatrix} \mathrm{I}_s & 0 \\ 0 & 0 \end{bmatrix}$ avec $s \leqslant r$
Pour $n \in \mathbb{N}^*$, considérons les matrices :

$$B_n = \begin{bmatrix} \mathrm{I}_s & & \\ & \frac{1}{n}\mathrm{I}_{r-s} & \\ & & 0 \end{bmatrix}$$

puis $A_n = PB_nQ$.
Alors B_n est de rang r, donc A_n aussi ; $\lim\limits_{n\to+\infty} B_n = J_s$, donc par continuité de l'application $M \mapsto PMQ$ (linéaire en dimension finie), on a aussi $\lim\limits_{n\to+\infty} A_n = A$.

A est donc limite d'une suite de matrices de E_r donc appartient à $\overline{E_r}$, ce qui démontre l'inclusion $F_r \subset \overline{E_r}$.

En conclusion, l'adhérence de l'ensemble des matrices de rang r est l'ensemble des matrices de rang inférieur ou égal à r.

Exercice 7.13

1) Démontrons l'inégalité proposée par récurrence sur $n \in \mathbb{N}^*$.
• Le résultat est immédiat si $n = 1$.
• Supposons l'inégalité établie à l'ordre n.
On écrit alors :

$$\begin{aligned}(A+B)^{n+1} - B^{n+1} &= \big((A+B)^n - B^n\big)B + (A+B)^n A,\end{aligned}$$

d'où, en utilisant l'inégalité triangulaire, le fait que $\|\cdot\|$ est une norme d'algèbre et l'hypothèse de récurrence :

$$\begin{aligned}&\big\|(A+B)^{n+1} - B^{n+1}\big\| \\ &\quad\leqslant \|(A+B)^n - B^n\|\,\|B\| \\ &\qquad\qquad + \|(A+B)^n\|\,\|A\| \\ &\quad\leqslant n\,\|A\|\,\big(\|A\| + \|B\|\big)^{n-1}\,\|B\| \\ &\qquad\qquad + \|A+B\|^n\,\|A\| \\ &\quad\leqslant n\,\|A\|\big(\|A\| + \|B\|\big)^{n-1}\big(\|A\| + \|B\|\big) \\ &\qquad\qquad + \big(\|A\| + \|B\|\big)^n\,\|A\| \\ &\quad\leqslant (n+1)\,\|A\|\,\big(\|A\| + \|B\|\big)^n\,,\end{aligned}$$

ce qui est le résultat cherché.

2) Soient $A,\ B \in \mathcal{M}_p(\mathbb{K})$. L'inégalité démontrée ci-dessus permet d'écrire (toutes les séries considérées étant absolument convergentes) :

$$\begin{aligned}&\|\exp(A+B) - \exp(B)\| \\ &\qquad= \left\|\sum_{n=1}^{+\infty} \frac{1}{n!}\big((A+B)^n - B^n\big)\right\| \\ &\qquad\leqslant \sum_{n=1}^{+\infty} \frac{1}{n!} n\,\|A\|\,\big(\|A\| + \|B\|\big)^{n-1} \\ &\qquad\leqslant \|A\| \sum_{n=1}^{+\infty} \frac{\big(\|A\| + \|B\|\big)^{n-1}}{(n-1)!} \\ &\qquad\leqslant \|A\|\exp\big(\|A\| + \|B\|\big)\,.\end{aligned}$$

Cela montre que l'application exp sur $\mathcal{M}_p(\mathbb{K})$ est localement lipschitzienne ; elle est donc continue.

Exercice 7.14

1) • Supposons d'abord A et B inversibles. Alors $\widetilde{A} = (\det A)A^{-1}$ et $\widetilde{B} = (\det B)B^{-1}$ donc :

$$\begin{aligned}\widetilde{B}\widetilde{A} &= (\det A.\det B)B^{-1}A^{-1} \\ &= \det(AB)(AB)^{-1} = \widetilde{AB}\,.\end{aligned}$$

• L'application qui à une matrice M associe sa comatrice est continue sur $\mathcal{M}_n(\mathbb{K})$, puisque ses applications coordonnées sont données à l'aide de déterminants extraits de M, et ces déterminants extraits sont des fonctions polynomiales donc continues en les coefficients de M.

L'application qui à tout couple de matrices (M,N) associe leur produit MN est continue sur $\mathcal{M}_n(\mathbb{K})^2$, car bilinéaire sur un espace vectoriel de dimension finie.
Enfin, l'ouvert $\mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})^2$, puisque l'on sait que toute matrice de $\mathcal{M}_n(\mathbb{K})$ est limite d'une suite de matrices inversibles (voir **[SF7.34]** page 394).

D'après la question précédente, les applications $(A,B) \mapsto \widetilde{AB}$ et $(A,B) \mapsto \widetilde{B}\,\widetilde{A}$ coïncident sur $\mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$, et elles sont continues sur $\mathcal{M}_n(\mathbb{K})^2$ comme composées d'applications continues.
Deux applications continues qui sont égales sur une partie dense sont égales partout ; cela donne le résultat demandé.

2) • Dans le cas D inversible, on écrit le produit par blocs :

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \times \begin{bmatrix} D & 0 \\ -C & D^{-1} \end{bmatrix} = \begin{bmatrix} AD - BC & BD^{-1} \\ 0 & \mathrm{I}_n \end{bmatrix},$$

et le calcul des déterminants par blocs donne le résultat demandé.

• Dans le cas général, on remarque d'abord que les applications $D \mapsto \det(AD - BC)$ et $D \mapsto \det \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ sont continues, car ce sont des fonctions polynomiales des coefficients de D.
Si D est une matrice quelconque de $\mathcal{M}_n(\mathbb{K})$, on a démontré en **[SF7.34]** qu'il existe une suite de matrices inversibles (D_p) qui converge vers D et qui sont de la forme $D + \lambda_p \mathrm{I}_n$. Donc si $CD = DC$ on a aussi $D_pC = CD_p$ pour tout $p \in \mathbb{N}$; on peut donc appliquer le résultat précédent aux matrices D_p, puis on obtient le résultat voulu par passage à la limite.

Exercice 8.1

Déjà on montre facilement par récurrence que $u_n \in \left[\frac{3}{4}\,;\frac{5}{4}\right]$ pour $n \geqslant 1$.
Soit $f\colon x \mapsto 1 + \frac{1}{4}\sin\frac{1}{x}$ pour $x \in \left[\frac{3}{4}\,;\frac{5}{4}\right]$.
Cet intervalle est stable par f et sur cet intervalle on a :

$$\left|f'(x)\right| = \left|-\frac{1}{4x^2}\cos\frac{1}{x}\right| \leqslant \frac{1}{4x^2} \leqslant \frac{4}{9}\cdot$$

D'après l'inégalité des accroissements finis on en tire :

$$|f(x) - f(y)| \leqslant \frac{4}{9}\,|x - y| \quad (*)$$

pour tous x, $y \in \left[\frac{3}{4}\,;\frac{5}{4}\right]$.

Ainsi f est contractante sur $\left[\frac{3}{4}\,;\frac{5}{4}\right]$. On applique alors la méthode vue en **[SF8.7]**.

• Soit $g\colon x \mapsto f(x) - x$.
Alors $g'(x) = f'(x) - 1 < 0$ donc g est strictement décroissante sur $\left[\frac{3}{4}\,;\frac{5}{4}\right]$.
Puisque $g\left(\frac{3}{4}\right) = \frac{1}{4} + \sin\frac{4}{3} > 0$ et $g\left(\frac{5}{4}\right) = -\frac{1}{4} + \frac{1}{4}\sin\frac{4}{5} < 0$, g s'annule une fois et une seule sur $\left[\frac{3}{4}\,;\frac{5}{4}\right]$ (théorème de la bijection, g étant continue).

Il existe donc un et un seul $\ell \in \left[\frac{3}{4}\,;\frac{5}{4}\right]$ tel que $f(\ell) = \ell$.

• D'après $(*)$ on aura, pour tout $n \geqslant 2$:

$$|u_n - \ell| = |f(u_{n-1}) - f(\ell)| \leqslant \frac{4}{9}\,|u_{n-1} - \ell|$$

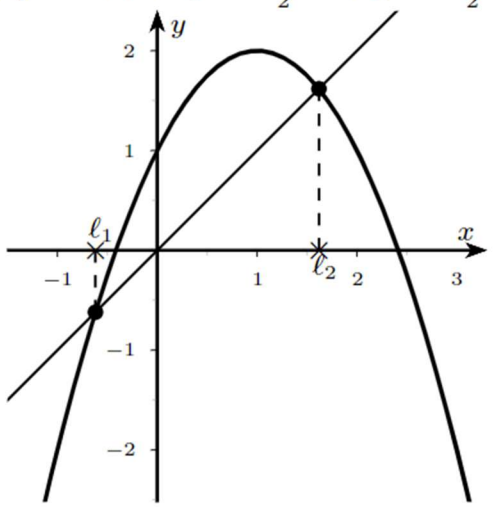
donc par récurrence :

$$\forall\, n \geqslant 1,\ |u_n - \ell| \leqslant \left(\frac{4}{9}\right)^{n-1} |u_1 - \ell|\ .$$

On en déduit : $\lim\limits_{n\to+\infty} u_n = \ell$.

Exercice 8.2

1) L'étude (non reproduite) et le tracé (ci-dessous) ne doivent pas poser de problème.
La résolution de l'équation du second degré $f(x) = x$ donne $\ell_1 = \frac{1-\sqrt{5}}{2}$ et $\ell_2 = \frac{1+\sqrt{5}}{2}$.



2) • L'étude précédente montre que $f(]-\infty\,;\ell_1[) =]-\infty\,;\ell_1[$, c'est-à-dire que $]-\infty\,;\ell_1[$ est stable par f.

Par conséquent, si $u_0 \in]-\infty\,;\ell_1[$, on aura $u_n \in]-\infty\,;\ell_1[$ pour tout n.

• Pour $x < \ell_1$, on a $f(x) < x$ (étudier le signe d'un trinôme) donc, puisque $u_n < \ell_1$, on aura $f(u_n) < u_n$ soit $u_{n+1} < u_n$: la suite est strictement décroissante.

• Si (u_n) était convergente, ce serait vers un réel ℓ tel que $f(\ell) = \ell$ puisque f est continue (en passant à la limite dans l'égalité $u_{n+1} = f(u_n)$).

De plus, la suite étant décroissante, on devrait avoir $\ell \leqslant u_n$ pour tout n, et en particulier $\ell \leqslant u_0 < \ell_1$. Or il n'existe pas de point fixe de f qui est strictement inférieur à ℓ_1, d'où la contradiction.

(u_n) étant décroissante non convergente, on a, d'après le cours : $\lim\limits_{n\to+\infty} u_n = -\infty$.

3) L'intervalle $]1\,;2[$ est stable par f et f est décroissante sur cet intervalle.
On sait donc déjà d'après le cours que les deux suites extraites (u_{2n}) et (u_{2n+1}) sont monotones de sens de variations contraires.

• Pour tout $x \in \mathbb{R}$:

$$\begin{aligned} &f\circ f(x) - x \\ &= -(-x^2+2x+1)^2 \\ &\quad + 2(-x^2+2x+1) + 1 - x \\ &= -(x^4 + 4x^2 + 1 - 4x^3 - 2x^2 + 4x) \\ &\quad - 2x^2 + 4x + 2 + 1 - x \\ &= -x^4 + 4x^3 - 4x^2 - x + 2 \\ &= (x-1)(x-2)(-x^2+2x+1) \\ &= -(x-1)(x-2)(x-\ell_1)(x-\ell_2)\,. \end{aligned}$$

Les 4 racines sont simples, donc le polynôme change de signe à chaque racine, et puisque $\lim\limits_{x\to\pm\infty} f\circ f(x) - x = -\infty$, on a facilement le tableau de signes :

x	$-\infty$		ℓ_1		1		ℓ_2		2		$+\infty$
$f\circ f(x)-x$		$-$	0	$+$	0	$-$	0	$+$	0	$-$	

• Puisque

$$f(]1\,;\ell_2[) =]\ell_2\,;2[\text{ et } f(]\ell_2\,;2[) =]1\,;\ell_2[\,,$$

on aura $f\circ f(]1\,;\ell_2[) =]1\,;\ell_2[$ et $f\circ f(]\ell_2\,;2[) =]\ell_2\,;2[$.

Ainsi les intervalles $]1\,;\ell_2[$ et $]\ell_2\,;2[$ sont des intervalles stables par $f\circ f$.

– Si $u_0 = \ell_2$, la suite u est constante.

– Si $u_0 \in]1\,;\ell_2[$, $u_2 - u_0 = f\circ f(u_0) - u_0$ est négatif, d'après le tableau de signes précédent.
Donc la suite (u_{2n}) est décroissante,et la suite (u_{2n+1}) est croissante.

Puisque (u_{2n}) reste dans l'intervalle $]1\,;\ell_2[$, elle est minorée par 1, donc elle converge ; sa limite ℓ doit vérifier $f\circ f(\ell) = \ell$ puisque $f\circ f$ est continue, doit appartenir à $[1\,;\ell_2]$ et doit être inférieure à u_0, donc est différente de ℓ_2 ; cette limite ne peut donc être que 1.

Puis, par continuité de f :

$$\lim_{n\to+\infty} u_{2n+1} = \lim_{n\to+\infty} f(u_{2n}) = f(1) = 2.$$

Puisqu'il existe deux suites extraites de (u_n) qui ont des limites différentes, la suite (u_n) diverge.

– Lorsque $u_0 \in]\ell_2\,;2[$, alors $u_1 \in]1\,;\ell_2[$ et on est ramené au cas précédent.

Exercice 8.3

• Il est facile de vérifier par récurrence que la suite $(y_n)_{n\in\mathbb{N}}$ est à termes strictement positifs, ce qui prouve aussi que la définition de l'énoncé a bien un sens.

• Pour $n \in \mathbb{N}^*$ on calcule :

$$\begin{aligned} y_n - 1 &= \frac{x_n + y_{n-1} - 1 - x_n y_{n-1}}{1 + x_n y_{n-1}} \\ &= \frac{(x_n - 1)(1 - y_{n-1})}{1 + x_n y_{n-1}}\,, \end{aligned}$$

ce qui permet de démontrer immédiatement par récurrence que :

$$\forall n \in \mathbb{N},\ 0 < y_n < 1\,.$$

• On aura alors pour tout $n \in \mathbb{N}^*$:

$$y_n - y_{n-1} = \frac{x_n(1 - y_{n-1}^2)}{1 + x_n y_{n-1}} > 0$$

donc la suite $(y_n)_{n\in\mathbb{N}}$ est croissante ; Étant majorée, elle converge vers un réel $\ell \in [0\,;1]$.

• La relation de l'énoncé peut aussi s'écrire :

$$x_n(y_n y_{n-1} - 1) = y_{n-1} - y_n\,,$$

donc puisque $\lim\limits_{n\to+\infty}(y_{n-1}-y_n)=0$ et que $\lim\limits_{n\to+\infty}(y_n y_{n-1}-1)=\ell^2-1$, si on avait $\ell\neq 1$ la suite $(x_n)_{n\in\mathbb{N}}$ convergerait vers 0, ce qui n'est pas possible.

On a donc prouvé par l'absurde que $\lim\limits_{n\to+\infty} y_n=1$.

Exercice 8.4

• On démontre par une récurrence immédiate que u_n et v_n sont strictement positifs pour tout n, ce qui implique dans le même temps que les deux suites sont bien définies (le dénominateur u_n+v_n ne peut s'annuler).

• Le calcul de $u_{n+1}-u_n$ montre que, pour connaître le sens de variation des suites, il faut d'abord connaître la position de u_n par rapport à v_n.

On calcule donc :

$$u_{n+1}-v_{n+1}=\frac{(u_n+v_n)^2-4u_nv_n}{2(u_n+v_n)}=\frac{(u_n-v_n)^2}{2(u_n+v_n)}\,,$$

d'où $u_{n+1}-v_{n+1}\geqslant 0$. On a donc $u_n\geqslant v_n$ pour $n\geqslant 1$.

• On en déduit, pour $n\geqslant 1$:

$$u_{n+1}-u_n=\frac{v_n-u_n}{2}\leqslant 0$$

donc la suite u est décroissante, au moins à partir du rang 1.

• Puis : $\dfrac{v_{n+1}}{v_n}=\dfrac{2u_n}{u_n+v_n}\geqslant 1$ pour $n\geqslant 1$: la suite v est croissante, au moins à partir du rang 1.

✓ $\dfrac{v_{n+1}}{v_n}\geqslant 1\implies v_{n+1}\geqslant v_n$ n'est vrai que parce que la suite v est à termes strictement positifs !

• On a donc, pour $n\geqslant 1$:

$$0<v_1\leqslant\ldots\leqslant v_n\leqslant u_n\leqslant\ldots\leqslant u_1\,.$$

La suite $(u_n)_{n\in\mathbb{N}^*}$ étant décroissante et minorée par v_1 converge, vers un certain réel ℓ ; la suite $(v_n)_{n\in\mathbb{N}^*}$ étant croissante et majorée par u_1 converge, vers un certain réel ℓ'.

Par passage à la limite dans les relations de récurrence, on a :

$$\ell=\frac{\ell+\ell'}{2}\quad\text{et}\quad\ell'=\frac{2\ell\ell'}{\ell+\ell'}\,,$$

donc $\ell=\ell'$ (et les suites sont adjacentes).

• Enfin, on remarque que $u_{n+1}v_{n+1}=u_nv_n$ donc la suite (u_nv_n) est constante, égale à u_0v_0. Par passage à la limite on a $\ell^2=u_0v_0$ donc $\ell=\sqrt{u_0v_0}$ (car $\ell\geqslant 0$).

Exercice 8.5

1) Le calcul :

$$x^3+y^3+z^3-3xyz=\frac{1}{2}(x+y+z)\times\left[(x-y)^2+(y-z)^2+(z-x)^2\right]$$

fournit directement le résultat demandé.

2) • On remarque d'abord que les définitions de l'énoncé ont bien un sens puisque par récurrence on peut montrer facilement que u_n, v_n et w_n sont strictement positifs pour tout n.

• $3w_{n+1}-3v_{n+1}\geqslant 0$ d'après la 1re question d'où $w_{n+1}\geqslant v_{n+1}$.

• $\dfrac{3}{u_{n+1}}-\dfrac{3}{v_{n+1}}\geqslant 0$ toujours d'après la 1re question d'où $v_{n+1}\geqslant u_{n+1}$.

Finalement :

$$\forall\, n\in\mathbb{N},\ u_{n+1}\leqslant v_{n+1}\leqslant w_{n+1}\,.$$

• On en déduit que, pour tout $n\in\mathbb{N}$:

$$\frac{3}{u_{n+2}}=\frac{1}{u_{n+1}}+\frac{1}{v_{n+1}}+\frac{1}{w_{n+1}}\leqslant\frac{3}{u_{n+1}}\,,$$

donc $(u_n)_{n\in\mathbb{N}^*}$ est croissante.

On a aussi :

$$w_{n+2}-w_{n+1}=\frac{(u_{n+1}-w_{n+1})+(v_{n+1}-w_{n+1})}{3}\leqslant 0,$$

et $(w_n)_{n\in\mathbb{N}^*}$ est décroissante.

Enfin, la suite $(u_n)_{n\in\mathbb{N}^*}$ est majorée par w_1 et la suite $(w_n)_{n\in\mathbb{N}^*}$ est minorée par u_1, donc ces deux suites convergent.

• Enfin, pour $n\geqslant 1$:

$$\begin{aligned}w_{n+1}-u_{n+1}&=\frac{u_n+v_n+w_n}{3}-u_{n+1}\\&\leqslant\frac{u_n+v_n+w_n}{3}-u_n\\&\leqslant\frac{(v_n-u_n)+(w_n-u_n)}{3}\\&\leqslant\frac{2}{3}(w_n-u_n)\,,\end{aligned}$$

d'où l'on tire par récurrence :

$$\forall n \geqslant 1,\ 0 \leqslant w_n - u_n \leqslant \left(\frac{2}{3}\right)^{n-1} (w_1 - u_1)$$

puis $\lim\limits_{n\to+\infty} u_n = \lim\limits_{n\to+\infty} w_n$ et finalement les 3 suites convergent vers la même limite grâce à l'encadrement :

$$\forall n \in \mathbb{N},\ u_n \leqslant v_n \leqslant w_n\,.$$

Exercice 8.6

1) Limitons-nous au cas $\ell \in \mathbb{R}$; dans le cas $\ell = \pm\infty$, le principe est identique, mais la définition de la limite change.

Soit $\varepsilon > 0$. Par définition de la limite :

$$\exists n_0 \in \mathbb{N},\ n \geqslant n_0 \implies \ell - \varepsilon \leqslant \frac{u_{n+1}}{u_n} \leqslant \ell + \epsilon$$

donc en multipliant ces inégalités (tout est positif) on aura pour $n \geqslant n_0$:

$$(\ell - \varepsilon)^{n-n_0} u_{n_0} \leqslant u_n \leqslant (\ell + \varepsilon)^{n-n_0} u_{n_0}$$

d'où :

$$(\ell - \varepsilon)^{1-\frac{n_0}{n}} \sqrt[n]{u_{n_0}} \leqslant \sqrt[n]{u_n} \leqslant (\ell + \varepsilon)^{1-\frac{n_0}{n}} \sqrt[n]{u_{n_0}}\,.$$

Or $\lim\limits_{n\to+\infty} (\ell - \varepsilon)^{1-\frac{n_0}{n}} \sqrt[n]{u_{n_0}} = \ell - \varepsilon$ et $\lim\limits_{n\to+\infty} (\ell+\varepsilon)^{1-\frac{n_0}{n}} \sqrt[n]{u_{n_0}} = \ell+\varepsilon$ donc à partir d'un certain rang on aura :

$$\ell - 2\varepsilon \leqslant \sqrt[n]{u_n} \leqslant \ell + 2\varepsilon\,,$$

ce qui prouve que $\lim\limits_{n\to+\infty} \sqrt[n]{u_n} = \ell$.

2) Il s'agit ici d'une simple vérification puisque l'énoncé donne le contre-exemple.

Pour tout $n \in \mathbb{N}$, on a :

$$\frac{u_{2n+1}}{u_{2n}} = b \quad \text{et} \quad \frac{u_{2n+2}}{u_{2n+1}} = a\,,$$

donc la suite de terme général $\frac{u_{n+1}}{u_n}$ diverge puisqu'on vient d'en trouver deux suites extraites ayant des limites différentes.
Cependant, en considérant les deux suites extraites d'indices pairs et impairs, on a facilement :

$$\lim_{n\to+\infty} \sqrt[n]{u_n} = \sqrt{ab}\,.$$

Exercice 8.7

1) La fonction f_n est de classe $\mathscr{C}^\infty$ sur $[0\,;1]$, et $f_n'(x) = n(x^{n-1} - 1)$ est strictement négatif sur $[0\,;1[$. La fonction f_n est strictement décroissante :

x	0		x_n		1
$f_n(x)$	1	↘	0	↘	$2-n$

f_n étant continue strictement décroissante, elle réalise une bijection de $[0\,;1]$ sur $[2-n\,;1]$. Puisque $n \geqslant 2$, 0 appartient à $[2-n\,;1]$ donc possède un unique antécédent x_n par f_n.

Conclusion : il existe un et un seul $x_n \in [0\,;1]$ tel que $f_n(x_n) = 0$.

2) a) Nous allons étudier les variations de la suite (x_n).

✓ La méthode dans ce genre d'exercice est classique : puisqu'on connaît les variations de f_n, pour positionner x_{n+1} par rapport à x_n, il suffit de connaître le signe de $f_n(x_{n+1})$ (ou de $f_{n+1}(x_n)$). Pour cela, on peut étudier le signe de $f_{n+1} - f_n$.

Pour $x \in [0\,;1]$,

$$f_{n+1}(x) - f_n(x) = x^n(x-1) - x\,,$$

donc $f_{n+1}(x) - f_n(x) \leqslant 0$.

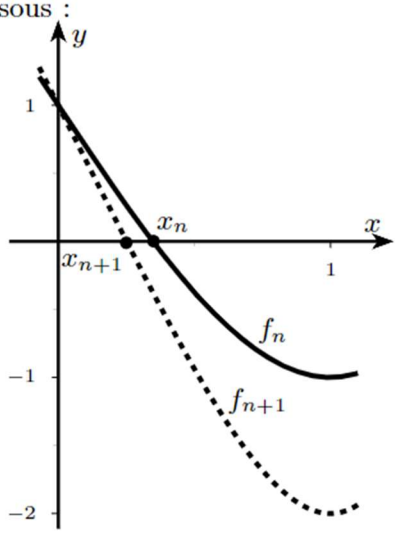
On aura donc en particulier

$$f_n(x_{n+1}) = f_n(x_{n+1}) - f_{n+1}(x_{n+1}) \geqslant 0\,,$$

et, compte tenu du tableau de variations de f_n, on en déduit que $x_{n+1} \leqslant x_n$.

Conclusion : la suite (x_n) est décroissante ; étant minorée par 0, elle converge.

Le principe utilisé est illustré par la figure ci-dessous :



b) Soit $a \in\ ]0\,;1[$. $\lim\limits_{n\to+\infty} a^n = 0$ donc

$$\lim_{n\to+\infty} f_n(a) = \lim_{n\to+\infty} (a^n - na + 1) = -\infty\,.$$

Il en résulte qu'à partir d'un certain rang, $f_n(a) < 0$, donc, compte tenu du tableau de variations de f_n, on aura $x_n < a$.

On a ainsi a démontré :

$$\forall\, a \in\ ]0\,;1[,\ \exists\, n_0 \in \mathbb{N} \text{ tel que } n \geqslant n_0 \implies 0 < x_n < a\,.$$

En remplaçant le a par un ε, on reconnaît la définition de la limite ! Donc :

$$\lim_{n\to+\infty} x_n = 0\,.$$

✓ Ce résultat pouvait s'obtenir par une autre méthode que celle suggérée par l'énoncé. En effet, la suite étant décroissante, on aura, pour $n \geqslant 3$, $0 < x_n \leqslant x_3 < 1$, d'où $0 < x_n^n \leqslant x_3^n$ et $\lim\limits_{n\to\infty} x_n^n = 0$. Puisque $x_n = \dfrac{x_n^n + 1}{n}$, on en déduit directement $\lim\limits_{n\to+\infty} x_n = 0$.

3) a) D'après la remarque ci-dessus, $\lim\limits_{n\to+\infty} x_n^n = 0$. Or $nx_n = x_n^n + 1$. Donc $\lim\limits_{n\to+\infty} nx_n = 1$ c'est-à-dire :

$$x_n \underset{n\to+\infty}{\sim} \frac{1}{n}\,.$$

b) L'équivalent obtenu ci-dessus s'écrit aussi $x_n = \dfrac{1}{n} + y_n$, où $y_n = o\left(\dfrac{1}{n}\right)$.

La question posée revient donc à chercher un équivalent de y_n. La méthode consiste à remplacer x_n dans l'égalité $f_n(x_n) = 0$ par le développement limité que l'on vient d'obtenir afin de trouver une relation sur y_n et d'en déduire un équivalent.

On traduit donc la relation $x_n^n = nx_n - 1$ sous la forme :

$$\left(\frac{1}{n} + y_n\right)^n = n\left(\frac{1}{n} + y_n\right) - 1 = ny_n\,,$$

d'où :

$$ny_n = \mathrm{e}^{n\ln\left(\frac{1}{n}+y_n\right)} = \mathrm{e}^{n\left[\ln\frac{1}{n}+\ln(1+ny_n)\right]} = \mathrm{e}^{n\left[\ln\frac{1}{n}+ny_n+o(ny_n)\right]}\,,$$

le développement limité du ln étant possible puisque ny_n tend vers 0 quand $n \to +\infty$. Ainsi : $ny_n = \mathrm{e}^{-n\ln n}\mathrm{e}^{n(ny_n+o(ny_n))}$.

Mais

$$y_n = nx_n - 1 = x_n^n < x_3^n$$

pour $n \geqslant 3$, et $0 < x_3 < 1$, donc, par croissances comparées,

$$\lim_{n\to\infty} n(ny_n + o(ny_n)) = 0\,,$$

et finalement

$$ny_n \underset{n\to+\infty}{\sim} \mathrm{e}^{-n\ln n} = \frac{1}{n^n}\,,$$

d'où : $y_n \underset{n\to+\infty}{\sim} \dfrac{1}{n^{n+1}}$.

En conclusion :

$$x_n \underset{n\to+\infty}{=} \frac{1}{n} + \frac{1}{n^{n+1}} + o\left(\frac{1}{n^{n+1}}\right).$$

Exercice 8.8

1) Déjà, il est facile de montrer par récurrence que $u_n > 0$ pour tout n, ce qui prouve que la suite u est bien définie.

Cela montre aussi que cette suite est croissante, puisque $u_{n+1} - u_n = \dfrac{1}{u_n} > 0$.
Si elle était convergente, ce serait vers un réel ℓ tel que $\ell \geqslant u_0 > 0$ et $\ell = \ell + \dfrac{1}{\ell}$: c'est impossible.
On en déduit que la suite diverge, forcément vers $+\infty$ puisqu'elle est croissante.

2) • Nous pourrions ici employer la méthode générale vue en **[SF8.8]**, mais il est facile de remarquer directement que l'on a, pour tout $n \in \mathbb{N}$:

$$u_{n+1}^2 - u_n^2 = 2 + \frac{1}{u_n^2}\,,$$

donc $\lim\limits_{n\to+\infty}(u_{n+1}^2 - u_n^2) = 2$.
Puisque la série $\sum\limits_{n\in\mathbb{N}} 2$ est divergente à termes positifs, le théorème de sommation des relations de comparaison donne :

$$\sum_{k=0}^{n-1}(u_{k+1}^2 - u_k^2) \underset{n\to+\infty}{\sim} \sum_{k=0}^{n-1} 2 = 2n\,,$$

d'où après télescopage :

$$u_n^2 \underset{n\to+\infty}{\sim} 2n\,,$$

puis : $u_n \underset{n\to+\infty}{\sim} \sqrt{2n}$.

• Posons, pour tout $n \in \mathbb{N}$:

$$v_n = u_n^2 - 2\,.$$

On a alors :

$$v_{n+1} - v_n = (u_{n+1}^2 - u_n^2) - 2 = \frac{1}{u_n^2} \underset{n\to+\infty}{\sim} \frac{1}{2n}\,.$$

Puisque la série $\sum_{n\in\mathbb{N}} \frac{1}{n}$ est divergente à termes positifs, le théorème de sommation des relations de comparaison donne :

$$\sum_{k=1}^{n-1}(v_{k+1}-v_k) \underset{n\to+\infty}{\sim} \sum_{k=1}^{n-1}\frac{1}{2k} \underset{n\to+\infty}{\sim} \frac{1}{2}\ln n\,,$$

d'où après télescopage :

$$v_n - v_1 \underset{n\to+\infty}{\sim} \frac{1}{2}\ln n\,.$$

Il en résulte $\lim_{n\to+\infty} v_n = +\infty$ donc $v_n \underset{n\to+\infty}{\sim} \frac{1}{2}\ln n$ soit :

$$v_n = \frac{1}{2}\ln n + o(\ln n)\ .$$

Il ne reste plus qu'à faire un développement limité :

$$\begin{aligned} u_n &= \sqrt{2n+v_n} = \left(2n+\frac{1}{2}\ln n + o(\ln n)\right)^{1/2} \\ &= \sqrt{2n}\left(1+\frac{1}{4}\frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)\right)^{1/2} \\ &= \sqrt{2n}\left(1+\frac{1}{8}\frac{\ln n}{n} + o\left(\frac{\ln n}{n}\right)\right) \\ &= \sqrt{2n} + \frac{\sqrt{2}}{8}\frac{\ln n}{\sqrt{n}} + o\left(\frac{\ln n}{\sqrt{n}}\right), \end{aligned}$$

ce qui est le développement asymptotique cherché.

Exercice 8.9

• L'implication :

$$\lim_{n\to+\infty} x_n = 0 \implies \lim_{n\to+\infty} y_n = 0$$

est claire.

• Supposons donc $\lim_{n\to+\infty} y_n = 0$. On commence par démontrer (récurrence par exemple) que :

$$\begin{aligned} x_n &= y_n + ky_{n-1} + \cdots + k^{n-1}y_1 + k^n y_0 \\ &= \sum_{i=0}^{n} k^{n-i}y_i \quad (*) \end{aligned}$$

puis on s'inspire de la démonstration du théorème de Cesàro.

Soit $\varepsilon > 0$. Par définition de la limite :

$$\exists\, n_0,\ n \geqslant n_0 \implies |y_n| < \epsilon$$

donc en utilisant l'inégalité triangulaire on aura d'après $(*)$, pour $n \geqslant n_0$:

$$|x_n| \leqslant \sum_{i=0}^{n_0-1} k^{n-i}\,|y_i| + \varepsilon \sum_{i=n_0}^{n} k^{n-i} \quad (**)$$

Or $\sum_{i=n_0}^{n} k^{n-i} \leqslant \sum_{i=0}^{+\infty} k^i = \frac{1}{1-k}$ (série géométrique) ; de plus, la suite $(y_n)_{n\in\mathbb{N}}$ étant convergente est bornée donc d'après $(**)$:

$$|x_n| \leqslant \|y\|_\infty \sum_{i=0}^{n_0-1} k^{n-i} + \frac{\varepsilon}{1-k}\,.$$

Puisque n_0 est fixé et que $k \in\ ]0\,;1[$, $\lim_{n\to+\infty} \sum_{i=0}^{n_0-1} k^{n-i} = 0$ donc cette quantité sera $< \varepsilon$ pour n assez grand.Ainsi :

$$|x_n| \leqslant \|y\|_\infty\, \varepsilon + \frac{\varepsilon}{1-k}$$

à partir d'un certain rang, ce qui prouve que $\lim_{n\to+\infty} x_n = 0$.

Exercice 8.10

• Posons $y_n = x_n - \frac{2}{3}$. Alors

$$x_n + \frac{x_{2n}}{2} = y_n + \frac{y_{2n}}{2} + 1\,,$$

donc $\lim_{n\to+\infty}\left(y_n + \frac{y_{2n}}{2}\right) = 0$ et on doit montrer que $\lim_{n\to+\infty} y_n = 0$.

• On montre déjà que toute suite extraite de la suite $(y_n)_{n\in\mathbb{N}}$ qui converge converge nécessairement vers 0. En effet, si $(y_{\varphi(n)})$ est une telle suite, si elle convergeait vers $\ell \neq 0$, la suite $(y_{2\varphi(n)})$ convergerait vers -2ℓ (car $y_{2\varphi(n)} + 2y_{\varphi(n)}$ tend vers 0), puis de même la suite $(y_{4\varphi(n)})$ convergerait vers 4ℓ etc. . . la suite $(y_{2^N\varphi(n)})$ convergerait vers $\pm 2^N\ell$; N étant aussi grand que l'on veut et ℓ étant supposé non nul, on pourrait extraire de y une suite non bornée, ce qui est exclu.

• La suite y est donc une suite bornée (car la suite x l'est) qui n'admet qu'une seule valeur d'adhérence ; d'après le résultat démontré en **[SF7.5]** page 372, elle converge vers cette valeur, c'est-à-dire vers 0.

Exercice 9.1

• Le résultat est immédiat si x est égal à l'un des a_i ; c peut alors prendre n'importe quelle valeur.

• Supposons donc $x \notin \{a_1, \dots, a_n\}$. Il existe alors un réel λ tel que :

$$f(x) = \lambda \frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{n!}\,.$$

Considérons la fonction g qui à tout $t \in [a\,;b]$ associe :

$$g(t) = f(t) - \lambda \frac{(t-a_1)(t-a_2)\cdots(t-a_n)}{n!}\,.$$

Alors g est de classe $\mathscr{C}^n$, comme f. g s'annule en $n+1$ points distincts (les a_i et x) donc d'après le théorème de Rolle g' s'annule en n points distincts puis g'' s'annule en $n-1$ points distincts etc...

Finalement, il existe $c \in [a\,;b]$ tel que $g^{(n)}(c) = 0$. Or le polynôme :

$$\frac{(X-a_1)(X-a_2)\cdots(X-a_n)}{n!}$$

est un polynôme de degré n dont le coefficient dominant est $\dfrac{X^n}{n!}$ donc sa dérivée n^{e} est égale à 1 et $g^{(n)}(t) = f^{(n)}(t) - \lambda$.

La relation $g^{(n)}(c) = 0$ implique $\lambda = f^{(n)}(c)$ et en remplaçant on trouve bien :

$$f(x) = \frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{n!} f^{(n)}(c)\,.$$

Exercice 9.2

Soit f comme dans l'énoncé, et soit $(a_n)_{n\in\mathbb{N}}$ une suite de valeurs deux à deux distinctes de $[a\,;b]$ où f s'annule.
D'après le théorème de Bolzano-Weierstrass, on peut extraire de cette suite une suite convergente $\big(a_{\varphi(n)}\big)$. Soit ℓ sa limite. Par continuité de f, les relations $f\big(a_{\varphi(n)}\big) = 0$ impliquent $f(\ell) = 0$.

De plus, en appliquant le théorème de Rolle entre $a_{\varphi(n)}$ et $a_{\varphi(n+1)}$ (qui sont distincts), il existe b_n strictement compris entre ces deux nombres tel que $f'(b_n) = 0$.
Par encadrement, la suite (b_n) tend aussi vers ℓ, et les relations $f'(b_n) = 0$ donnent par passage à la limite puisque f' est continue : $f'(\ell) = 0$.

Puisque $f(\ell) = f'(\ell) = 0$, le théorème de Cauchy-Lipschitz implique que f est la fonction nulle sur $[a\,;b]$.

Exercice 9.3

1) • Puisque f est à valeurs positives, l'inégalité $f'' \geqslant af$ implique que f est convexe.
Si f' prenait une valeur strictement positive en un point x_0, puisque la courbe de f est située au dessus de sa tangente, on aurait :

$$\forall\, x \geqslant 0,\ f(x) \geqslant f'(x_0)(x - x_0) + f(x_0)\,,$$

ce qui impliquerait $\lim\limits_{x\to+\infty} f(x) = +\infty$. Cela est impossible puisque l'énoncé suppose f majorée.
On a donc montré par l'absurde que $f' \leqslant 0$: f est décroissante.

• f étant décroissante positive, elle admet en $+\infty$ une limite $\ell \geqslant 0$ d'après le théorème de la limite monotone. Supposons $\ell > 0$. On aurait donc $f''(t) \geqslant af(t) \geqslant a\ell > 0$ pour tout t d'où en intégrant entre 0 et x :

$$f'(x) - f'(0) \geqslant a\ell x\,.$$

Cela impliquerait $\lim\limits_{x\to+\infty} f'(x) = +\infty$, ce qui contredit l'inégalité $f' \leqslant 0$ trouvée auparavant.
On a donc montré par l'absurde que $\ell = 0$.

• De même f' est croissante majorée par 0, donc possède une limite ℓ' en $+\infty$. Supposons $\ell' < 0$. L'inégalité $f' \leqslant \ell$ donne alors, en intégrant :

$$f(x) - f(0) \leqslant \ell' x\,,$$

ce qui implique $\lim\limits_{x\to+\infty} f(x) = -\infty$, et contredit le résultat précédent.
On a donc montré par l'absurde que $\ell' = 0$.

2) Posons $g(x) = f(x)\mathrm{e}^{x\sqrt{a}}$ pour $x \geqslant 0$. Alors $g'(x) = \big(\sqrt{a}f(x) + f'(x)\big)\mathrm{e}^{x\sqrt{a}}$.
Posons ensuite :

$$\begin{aligned} h(x) &= g'(x)\mathrm{e}^{-2x\sqrt{a}} \\ &= \big(\sqrt{a}f(x) + f'(x)\big)\mathrm{e}^{-x\sqrt{a}}. \end{aligned}$$

On a $h'(x) = \big(-af(x) + f''(x)\big)\mathrm{e}^{-x\sqrt{a}}$ donc compte tenu de l'hypothèse, $h'(x) \geqslant 0$ pour tout $x \geqslant 0$.
On en déduit que h est croissante et, puisque $\lim\limits_{x\to+\infty} h(x) = 0$, on a $h(x) \leqslant 0$ sur $\mathbb{R}_+$.
On en déduit alors que g est décroissante sur $\mathbb{R}_+$, donc $g(x) \leqslant g(0)$ pour tout $x \geqslant 0$, ce qui donne l'inégalité demandée.

Exercice 9.4

Puisque $b \in]a\,;c[$, il existe $\lambda \in]0\,;1[$ tel que $b = \lambda a + (1-\lambda)c$.
f étant convexe, on a :

$$f(b) \leqslant \lambda f(a) + (1-\lambda) f(c) \quad (1)$$

D'autre part,

$$a - b + c = (1-\lambda)a + \lambda c\,,$$

donc :

$$f(a-b+c) \leqslant (1-\lambda) f(a) + \lambda f(c) \quad (2)$$

En additionnant (1) et (2) on obtient :

$$f(a-b+c) + f(b) \leqslant f(a) + f(c)\,,$$

ce qui est l'inégalité cherchée.

Exercice 9.5

Puisque l'on ne sait pas si g est dérivable, pour montrer que g est croissante, nous revenons simplement à la définition.
Soient donc x, $y \in \mathbb{R}_+$ tels que $x \leqslant y$. On calcule :

$$\begin{aligned} g(y) - g(x) &= y f'(y) - x f'(x) + f(x) - f(y) \\ &= (y-x) f'(y) + x\big(f'(y) - f'(x)\big) \\ &\qquad + f(x) - f(y)\,. \end{aligned}$$

D'après le théorème des accroissements finis, il existe $c \in]x\,;y[$ tel que

$$f(y) - f(x) = (y-x) f'(c)\,,$$

donc en remplaçant :

$$\begin{aligned} g(y) - g(x) = (y-x)\big(f'(y) - f'(c)\big) \\ + x\big(f'(y) - f'(x)\big)\,. \end{aligned}$$

Puisque f est convexe, f' est croissante donc $f'(y) - f'(c) \geqslant 0$, et aussi $f'(y) - f'(x) \geqslant 0$, donc $g(y) - g(x) \geqslant 0$ ce qui prouve que g est croissante.

Exercice 9.6

1) La fonction $F_c : t \mapsto \dfrac{f(t) - f(c)}{t - c}$ est croissante sur $[a\,;b] \setminus \{c\}$.
Or $F_c(a) = F_c(b) = 0$, donc F_c est nulle, c'est-à-dire que pour $t \neq c$, $f(t) = f(c)$. f est donc constante sur $[a\,;b]$.

2) Soit $\ell : x \mapsto \dfrac{f(b) - f(a)}{b-a}(x-a) + f(a)$ (ℓ est l'équation de la droite passant par les points de coordonnées $\big(a,f(a)\big)$ et $\big(b,f(b)\big)$), et soit $g = f - \ell$.
g est convexe, comme somme de deux fonctions convexes, et $g(a) = g(b) = g(c) = 0$ compte tenu de l'hypothèse.
D'après la question précédente, g est constante, donc g est nulle. Ainsi $f = \ell$ est affine.

Exercice 10.1

Notons $\overrightarrow{F}(t) = \overrightarrow{OM}(t) = x(t)\,\overrightarrow{i} + y(t)\,\overrightarrow{j}$.

• La fonction $\overrightarrow{F}$ est 2π-périodique ; on a donc $M(t + 2\pi) = M(t)$, et la courbe est entièrement obtenue lorsque t décrit un intervalle de longueur 2π, par exemple $t \in [-\pi\,;\pi]$.

• Le point $M(t + \pi)$ est le symétrique de $M(t)$ par rapport à O ; on peut donc se contenter d'étudier la courbe pour $t \in [0\,;\pi]$, puis on complètera par symétrie.

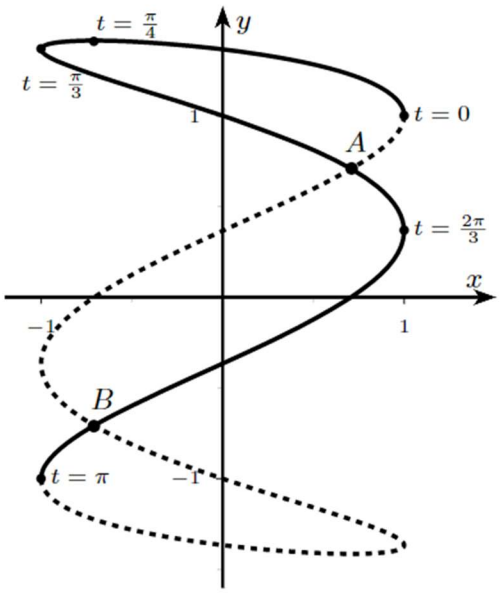
• Pour tout $t \in [0\,;\pi]$ on a

$$x'(t) = -3\sin 3t$$
$$y'(t) = -\sin t + \cos t = \sqrt{2}\cos\left(t + \frac{\pi}{4}\right).$$

Le tableau de variations est facile à former :

t	0		$\frac{\pi}{4}$		$\frac{\pi}{3}$		$\frac{2\pi}{3}$		π
x'	0	$-$		$-$	0	$+$	0	$-$	0
x	1	↘		↘	-1	↗	1	↘	-1
y	1	↗	$\sqrt{2}$	↘		↘		↘	-1
y'		$+$	0	$-$		$-$		$-$	

On obtient le tracé suivant. Nous avons tracé en trait plein la partie étudiée et en pointillés celle qui s'en déduit par symétrie par rapport à O.



Ce type de courbe s'appelle une *courbe de Lissajous.*

• Pour chercher les points doubles, il faut résoudre le système :

$$\begin{cases} \cos 3t &= \cos 3u \\ \cos t + \sin t &= \cos u + \sin u \end{cases}$$

avec $-\pi \leqslant t < u < \pi$.
La deuxième équation s'écrit aussi :

$$\sqrt{2}\cos\left(t - \frac{\pi}{4}\right) = \sqrt{2}\cos\left(u - \frac{\pi}{4}\right)$$

ce qui donne, compte tenu des contraintes sur t et u :

$$t - \frac{\pi}{4} \equiv \frac{\pi}{4} - u \pmod{2\pi}$$

soit

$$t \equiv \frac{\pi}{2} - u \pmod{2\pi}.$$

En remplaçant dans la 1re équation, on obtient :

$$\sin\left(\frac{3\pi}{2} - 3u\right) = \sin 3u$$

d'où la seule possibilité :

$$\frac{3\pi}{2} - 3u \equiv 3u \pmod{2\pi}$$

qui conduit à :

$$u \equiv \frac{\pi}{4} \pmod{\frac{\pi}{3}}.$$

$u = \frac{\pi}{4}$ donne $t = \frac{\pi}{4}$, ce qui est exclu. On essaie ainsi toutes les possibilités et on trouve comme seules solutions :

$$u = \frac{7\pi}{12} \quad \text{et} \quad t = -\frac{\pi}{12} \quad : \text{ point } A$$

et

$$u = \frac{11\pi}{12} \quad \text{et} \quad t = -\frac{5\pi}{12} \quad : \text{ point } B.$$

Exercice 10.2

1) Le plan étant rapporté à un repère orthonormé $(O; \vec{i}, \vec{j})$, on notera

$$\vec{F}(t) = \overrightarrow{OM}(t) = x(t)\vec{i} + y(t)\vec{j}.$$

• La fonction $t \mapsto \vec{F}(t)$ est, d'après les théorèmes du cours, de classe $\mathscr{C}^\infty$ sur chaque intervalle où la fonction sin ne s'annule pas, c'est-à-dire de la forme $]k\pi\,;(k+1)\pi[$ avec $k \in \mathbb{Z}$.

• Les fonctions x et y sont toutes deux π-périodiques ; ainsi $M(t) = M(t+\pi)$ et on obtiendra toute la courbe en restreignant l'étude à un intervalle de longueur π, par exemple $]0\,;\pi[$.

On remarque ensuite que, pour tout $t \in]0\,;\pi[$,

$$x(\pi - t) = x(t)$$
$$y(\pi - t) = -y(t)$$

donc les points $M(\pi - t)$ et $M(t)$ sont symétriques par rapport à l'axe Ox.
On peut donc limiter l'étude à $t \in \left]0\,;\frac{\pi}{2}\right]$ et l'on complètera la courbe obtenue par symétrie.

• On calcule alors, pour $t \in \left]0\,;\frac{\pi}{2}\right]$:

$$x'(t) = -2\cos t \sin t + \frac{\cos t}{\sin t}$$
$$= \frac{\cos t(1 - 2\sin^2 t)}{\sin t} = \frac{\cos t \cos(2t)}{\sin t}$$

et $y'(t) = \cos(2t)$.

Il y a ici quelque chose de remarquable : la pente de la tangente en un point régulier de paramètre t est $m(t) = \dfrac{y'(t)}{x'(t)} = \tan t$.

Le tableau de variations ne présente pas de difficulté particulière.

t	0		$\frac{\pi}{4}$		$\frac{\pi}{2}$
$x'(t)$	‖	$+$	0	$-$	0
$x(t)$	$-\infty$	↗	$\frac{1-\ln 2}{2}$	↘	0
$y(t)$	0	↗	$\frac{1}{2}$	↘	0
$y'(t)$		$+$	0	$-$	

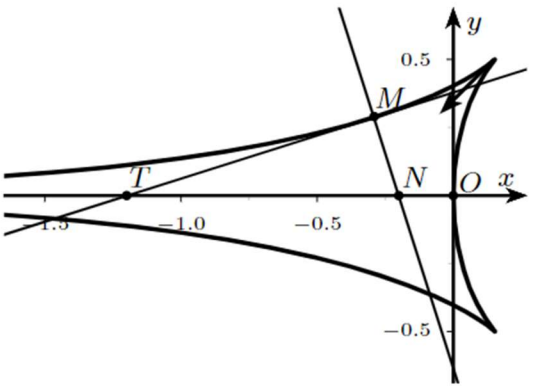
• Lorsque t tend vers 0^+, $x(t)$ tend vers $-\infty$ et $y(t)$ tend vers 0^+ : l'axe Ox est donc asymptote à la courbe, celle-ci étant située au-dessus.

• Pour $t = \frac{\pi}{2}$, le point M est en O, x' s'annule alors que y' ne s'annule pas : il y a une tangente verticale.

• Pour $t = \frac{\pi}{4}$, le vecteur $\overrightarrow{F}'(t)$ est nul : il s'agit donc d'un point stationnaire.
Pour préciser la pente de la tangente en ce point, il suffit de calculer

$$\lim_{t\to\frac{\pi}{4}} \frac{y'(t)}{x'(t)} = \lim_{t\to\frac{\pi}{4}} \tan t = 1\,.$$

• Il ne reste plus qu'à faire le tracé :



2) Supposons $t \in \left]0\,;\frac{\pi}{2}\right[\setminus \left\{\frac{\pi}{4}\right\}$. Puisque $x'(t) = \frac{\cos t\cos(2t)}{\sin t}$ et $y'(t) = \cos(2t)$, un vecteur tangent à la courbe en $M(t)$ est $\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}$. On obtient alors :

– l'équation de la tangente en $M(t)$:

$$\sin t\big(x - x(t)\big) - \cos t\big(y - y(t)\big) = 0\,,$$

– et celle de la normale :

$$\cos t\big(x - x(t)\big) + \sin t\big(y - y(t)\big) = 0\,.$$

Les abscisses de T et de N, notées respectivement x_T et x_N s'obtiennent en faisant $y = 0$ dans les équations correspondantes :

$$x_T = \frac{-y(t)\cos t + x(t)\sin t}{\sin t}$$
$$x_N = \frac{y(t)\sin t + x(t)\cos t}{\cos t}$$

d'où : $\overline{TN} = x_N - x_T$

$$= \frac{-y(t)(\cos^2 t + \sin^2 t)}{\sin t\cos t} = -1\,.$$

On a donc bien $TN = cste$

Exercice 10.3

1) Notons $I = \displaystyle\int_0^{\frac{\pi}{2}} \frac{\sin t}{1+\sin t+\cos t}\,.$
La fonction à intégrer est continue sur $\left[0\,;\frac{\pi}{2}\right]$ puisque le dénominateur ne s'annule pas, donc I existe.
Il n'y a pas ici d'astuce particulière ; on utilise le changement de variable standard $x = \tan\frac{t}{2}$. Puisque $t \in \left[0\,;\frac{\pi}{2}\right]$,

$$x = \tan\frac{t}{2} \iff t = 2\operatorname{Arctan} x\,,$$

donc, compte tenu de :

$$\sin t = \frac{2x}{1+x^2},\ \cos t = \frac{1-x^2}{1+x^2},\ \mathrm{d}t = \frac{2\,\mathrm{d}x}{1+x^2}\,,$$

on obtient :

$$I = \int_0^1 \frac{2x}{(1+x)(1+x^2)}\,\mathrm{d}x$$
(décomposition en éléments simples)
$$= \int_0^1 \left(-\frac{1}{1+x} + \frac{x}{1+x^2} + \frac{1}{1+x^2}\right)\mathrm{d}x$$
$$= \left[-\ln(1+x) + \frac{1}{2}\ln(1+x^2) + \operatorname{Arctan} x\right]_0^1$$
$$= \frac{\pi}{4} - \frac{1}{2}\ln 2\,.$$

2) Notons $I = \displaystyle\int_0^{\pi} \frac{\mathrm{d}t}{1+a^2\sin^2 t}\,.$
L'existence de cette intégrale ne pose pas de problème.
Les règles de Bioche nous incitent à poser $x = \tan t$, mais la fonction tan n'est pas définie sur $[0\,;\pi]$! Il est donc plus judicieux de poser $x = \cot t$; ce changement de variable est licite puisque la fonction cot réalise une bijection strictement décroissante de classe $\mathscr{C}^1$ de l'intervalle $]0\,;\pi[$ sur $\mathbb{R}$. L'intégrale généralisée qui apparaît est donc convergente. Il ne reste plus qu'à calculer :

$$I = \int_0^{\pi} \frac{1}{a^2 + \frac{1}{\sin^2 t}}\cdot\frac{\mathrm{d}t}{\sin^2 t}$$
$$= \int_{+\infty}^{-\infty} \frac{1}{a^2+1+x^2}(-\mathrm{d}x)$$
$$= \int_{-\infty}^{+\infty} \frac{\mathrm{d}x}{x^2+a^2+1}$$
$$= \frac{1}{\sqrt{a^2+1}}\left[\operatorname{Arctan}\left(\frac{x}{\sqrt{a^2+1}}\right)\right]_{-\infty}^{+\infty}$$
$$= \frac{\pi}{\sqrt{a^2+1}}\,.$$

3) Soit $I = \displaystyle\int_1^{+\infty} \frac{\sqrt{t^3+1}}{t^4}\,\mathrm{d}t$.

La fonction $t \mapsto \dfrac{\sqrt{t^3+1}}{t^4}$ est continue positive sur $[1\,;+\infty[$ et est équivalente en $+\infty$ à $t \mapsto \dfrac{1}{t^{5/2}}$ qui est intégrable au voisinage de $+\infty$, donc I existe.

Aucune des règles énoncées dans le **[SF12]** ne s'applique ici. On réfléchit donc et on pense à écrire :

$$\frac{\sqrt{t^3+1}}{t^4}\,\mathrm{d}t = \underbrace{\frac{\sqrt{t^3+1}}{t^6}}_{\text{fonction de } t^3} \underbrace{t^2\,\mathrm{d}t}_{=\frac{1}{3}\,\mathrm{d}(t^3)}.$$

On fait donc le changement de variable $u = t^3$, ce qui est licite puisque l'application $t \mapsto t^3$ réalise une bijection strictement croissante de classe $\mathscr{C}^1$ de l'intervalle $[1\,;+\infty[$ sur lui-même. On a alors plus simplement :

$$I = \frac{1}{3}\int_1^{+\infty} \frac{\sqrt{u+1}}{u^2}\,\mathrm{d}u\,.$$

On effectue ensuite le changement de variable (licite) $x = \sqrt{u+1}$, soit $u = x^2-1$, ce qui donne :

$$I = \frac{2}{3}\int_{\sqrt{2}}^{+\infty} \frac{x^2}{(x^2-1)^2}\,\mathrm{d}x\,.$$

Comme dans l'exemple traité en **[SF10.8]** page 437, on fait une intégration par parties pour abaisser les degrés avant de faire une décomposition en éléments simples :

$$\begin{aligned} I &= \frac{2}{3}\int_{\sqrt{2}}^{+\infty} \underbrace{x}_{u}\,\underbrace{\frac{x}{(x^2-1)^2}}_{v'}\,\mathrm{d}x \\ &= \frac{2}{3}\left\{\left[\frac{-x}{2(x^2-1)}\right]_{\sqrt{2}}^{+\infty} + \int_{\sqrt{2}}^{+\infty} \frac{\mathrm{d}x}{2(x^2-1)}\right\} \\ &= \frac{\sqrt{2}}{3} + \frac{1}{6}\int_{\sqrt{2}}^{+\infty}\left(\frac{1}{x-1} - \frac{1}{x+1}\right)\mathrm{d}x \\ &= \frac{\sqrt{2}}{3} + \frac{1}{6}\left[\ln\frac{x-1}{x+1}\right]_{\sqrt{2}}^{+\infty} \\ &= \frac{\sqrt{2}+\ln(\sqrt{2}+1)}{3}\,. \end{aligned}$$

4) Soit $I = \displaystyle\int_1^{+\infty} \frac{t\ln t}{(1+t^2)^2}\,\mathrm{d}t$.

La fonction $t \mapsto \dfrac{t\ln t}{(1+t^2)^2}$ est continue positive sur $[1\,;+\infty[$ et est majorée par $t \mapsto \dfrac{1}{t^2}$ qui est une fonction de Riemann intégrable au voisinage de $+\infty$, donc I existe.

Le calcul fait appel à deux techniques classiques : intégration par parties(qui est justifiée car le terme entre crochets a une valeur finie) puis décomposition en éléments simples.

$$\begin{aligned} I &= \int_1^{+\infty} \underbrace{\ln t}_{u}\,\underbrace{\frac{t}{(1+t^2)^2}}_{v'}\,\mathrm{d}t \\ &= \underbrace{\left[-\frac{\ln t}{2(1+t^2)}\right]_1^{+\infty}}_{=0} + \frac{1}{2}\int_1^{+\infty} \frac{\mathrm{d}t}{t(1+t^2)} \\ &= \frac{1}{2}\int_1^{+\infty}\left(\frac{1}{t} - \frac{t}{1+t^2}\right)\mathrm{d}t \\ &= \left[\frac{1}{2}\ln\left(\frac{t}{\sqrt{t^2+1}}\right)\right]_1^{+\infty} = \frac{1}{4}\ln 2\,. \end{aligned}$$

Exercice 10.4

La fonction à intégrer étant paire, on a

$$a_n = \frac{2}{\pi}\int_0^{\pi} \frac{\cos nx}{4\cos x - 5}\,\mathrm{d}x\,.$$

1) • On effectue le changement de variable $t = \tan\dfrac{x}{2}$, qui réalise une bijection strictement monotone de classe $\mathscr{C}^1$ de $[0\,;\pi[$ sur $\mathbb{R}_+$, ce qui assure la convergence des intégrales généralisées qui suivent.
On sait alors que :

$$\cos x = \frac{1-t^2}{1+t^2} \quad\text{et}\quad \mathrm{d}x = \frac{2}{1+t^2}\,\mathrm{d}t\,,$$

ce qui mène à :

$$\begin{aligned} a_0 &= \frac{2}{\pi}\int_0^{+\infty} \frac{\frac{2}{1+t^2}}{4\left(\frac{1-t^2}{1+t^2}\right)-5}\,\mathrm{d}t \\ &= -\frac{4}{\pi}\int_0^{+\infty} \frac{1}{1+9t^2}\,\mathrm{d}t\,. \end{aligned}$$

En effectuant ensuite le changement de variable $u = 3t$, on obtient :

$$\begin{aligned} a_0 &= -\frac{4}{\pi}\int_0^{+\infty} \frac{1}{1+u^2}\,\frac{\mathrm{d}u}{3} \\ &= -\frac{4}{3\pi}\underbrace{\int_0^{+\infty} \frac{1}{1+u^2}\,\mathrm{d}u}_{=\frac{\pi}{2}} = -\frac{2}{3} \end{aligned}$$

• Pour calculer a_1 on pourrait utiliser le même changement de variable, mais une petite astuce (classique) permet de simplifier les calculs.

Cela consiste à faire apparaître le dénominateur $4\cos x-5$ au numérateur, de la façon suivante :

$$\begin{aligned} a_1 &= \frac{2}{\pi}\int_0^\pi \frac{\cos x}{4\cos x-5}\,\mathrm{d}x \\ &= \frac{2}{\pi}\int_0^\pi \frac{\left(\cos x-\frac{5}{4}\right)+\frac{5}{4}}{4\cos x-5}\,\mathrm{d}x \\ &= \frac{2}{\pi}\int_0^\pi \frac{1}{4}\,\mathrm{d}x + \frac{5}{4}a_0 = \frac{1}{2}-\frac{5}{6} = -\frac{1}{3}\cdot \end{aligned}$$

2) Soit $n\in\mathbb{N}$. Les formules usuelles de trigonométrie nous donnent :

$$\cos(n+2)x+\cos nx = 2\cos x\cos(n+1)x.$$

On obtient donc :

$$a_{n+2}+a_n = \frac{4}{\pi}\int_0^\pi \frac{\cos x\cos(n+1)x}{4\cos x+5}\,\mathrm{d}x.$$

En utilisant la même astuce de calcul que précédemment on obtient :

$$\begin{aligned} &a_{n+2}+a_n \\ &= \frac{4}{\pi}\left(\int_0^\pi \frac{\frac{1}{4}\left[4\cos x-5\right]\cos(n+1)x}{4\cos x-5}\,\mathrm{d}x \right. \\ &\qquad \left. + \int_0^\pi \frac{\frac{5}{4}\cos(n+1)x}{4\cos x-5}\,\mathrm{d}x\right), \end{aligned}$$

soit :

$$\begin{aligned} a_{n+2}+a_n &= \frac{1}{\pi}\underbrace{\int_0^\pi \cos(n+1)x\,\mathrm{d}x}_{=0} \\ &\quad + \frac{5}{\pi}\int_0^\pi \frac{\cos(n+1)x}{4\cos x-5}\,\mathrm{d}x \\ &= \frac{5}{2}\,a_{n+1}\cdot \end{aligned}$$

On a donc la relation :

$$\forall\, n\in\mathbb{N},\ a_{n+2}-\frac{5}{2}\,a_{n+1}+a_n = 0\,.$$

3) Il s'agit d'une relation de récurrence linéaire du second ordre, dont l'équation caractéristique est $X^2-\frac{5}{2}X+1=0$.

Ses racines sont 2 et $\frac{1}{2}\cdot$

On sait alors d'après le cours (cf. **[S8.6]**) qu'il existe des réels α, β tels que

$$\forall\, n\in\mathbb{N},\ a_n = \alpha 2^n+\beta\frac{1}{2^n}\cdot$$

Les valeurs initiales $a_0=-\frac{2}{3}$ et $a_1=-\frac{1}{3}$ permettent de trouver $\alpha=0$ et $\beta=-\frac{2}{3}\cdot$

En conclusion : $\forall\, n\in\mathbb{N},\ a_n = -\dfrac{1}{3\cdot 2^{n-1}}\cdot$

Exercice 10.5

Posons $v_n=\ln(u_n)$ (cela a bien un sens). Ainsi :

$$v_n = \sum_{k=1}^n \ln\left(1+\frac{\sqrt{k(n-k)}}{n^2}\right).$$

Sous cette forme, v_n ressemble à une somme de Riemann. Pour s'y ramener exactement, nous allons utiliser un développement limité de la fonction ln.
Plus précisément, posons pour $x>0$:

$$\varphi(x) = \frac{\ln(1+x)-x}{x^2}\,.$$

La fonction φ est continue sur $\mathbb{R}_+^*$ et en utilisant les développements limités usuels, on a $\lim\limits_{x\to 0^+}\varphi(x) = -\frac{1}{2}\cdot$ On peut donc prolonger φ en une fonction continue sur $\mathbb{R}_+$ et, par suite, φ est bornées sur $[0\,;1]$.

Pour $k\in[\![1\,;n]\!]$, $k(n-k)\leqslant\frac{n^2}{4}$ donc

$$\frac{\sqrt{k(n-k)}}{n^2}\leqslant\frac{1}{2n}\leqslant 1\,,$$

et la relation

$$\ln(1+x) = x+x^2\varphi(x)$$

permet d'écrire :

$$v_n = \sum_{k=1}^n \frac{\sqrt{k(n-k)}}{n^2}+r_n$$

avec

$$\begin{aligned} r_n &= \sum_{k=1}^n \frac{k(n-k)}{n^4}\varphi\left(\frac{\sqrt{k(n-k)}}{n^2}\right) \\ &\leqslant \|\varphi\|_\infty^{[0;1]}\sum_{k=1}^n \frac{k(n-k)}{n^4} \\ &\leqslant \|\varphi\|_\infty^{[0;1]}\sum_{k=1}^n \frac{1}{4n^2} = \frac{\|\varphi\|_\infty^{[0;1]}}{4n}\cdot \end{aligned}$$

Ainsi, $\lim\limits_{n\to+\infty} r_n = 0$.

D'autre part,

$$\sum_{k=1}^n \frac{\sqrt{k(n-k)}}{n^2} = \frac{1}{n}\sum_{k=1}^n \sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}$$

est une somme de Riemann associée à la fonction $x\mapsto\sqrt{x(1-x)}$, continue sur $[0\,;1]$, donc converge vers l'intégrale de cette fonction lorsque $n\to+\infty$.

On en déduit :

$$\begin{aligned}\lim_{n\to+\infty} v_n &= \int_0^1 \sqrt{x(1-x)}\,\mathrm{d}x \\ &= \frac{1}{2}\int_0^{\frac{\pi}{2}} \sin^2(2t)\,\mathrm{d}t \\ &\qquad \text{(changement de variable } x=\sin^2 t\text{)} \\ &= \frac{1}{4}\int_0^{\frac{\pi}{2}} (1-\cos 4t)\,\mathrm{d}t = \frac{\pi}{8}\,,\end{aligned}$$

puis finalement : $\lim\limits_{n\to+\infty} u_n = \mathrm{e}^{\frac{\pi}{8}}$.

Exercice 10.6

1) • f est bien définie comme intégrale d'une fonction continue sur un segment (le dénominateur ne s'annule pas).

• Pour étudier la monotonie de f, nul besoin de calculer sa dérivée (dont le calcul nécessite les résultats des théorèmes sur les intégrales à paramètre). On revient simplement à la définition.

Pour $0 < x < y$ on a $\frac{1}{x} > \frac{1}{y}$ donc :

$$\begin{aligned}&f(x) - f(y) \\ &= \int_0^{\frac{1}{x}} \frac{\mathrm{d}t}{x+\sin^2 t} - \int_0^{\frac{1}{y}} \frac{\mathrm{d}t}{y+\sin^2 t} \\ &= \int_0^{\frac{1}{y}} \underbrace{\left(\frac{1}{x+\sin^2 t} - \frac{1}{y+\sin^2 t}\right)}_{>0}\mathrm{d}t \\ &\qquad + \underbrace{\int_{\frac{1}{y}}^{\frac{1}{x}} \frac{\mathrm{d}t}{x+\sin^2 t}}_{>0}\,,\end{aligned}$$

d'où $f(x) > f(y)$ et f est strictement décroissante.

2) Pour tout $x > 0$ et tout réel t, on a :

$$x \leqslant x + \sin^2 t \leqslant x+1\,,$$

donc :

$$\int_0^{\frac{1}{x}} \frac{\mathrm{d}t}{x+1} \leqslant f(x) \leqslant \int_0^{\frac{1}{x}} \frac{\mathrm{d}t}{x}$$

c'est-à-dire :

$$\frac{1}{x(x+1)} \leqslant f(x) \leqslant \frac{1}{x^2}\,.$$

Cela donne immédiatement :

$$\lim_{x\to 0^+} f(x) = +\infty\,,\quad \lim_{x\to+\infty} f(x) = 0\,,$$

et aussi l'équivalent (demandé à la question suivante) : $f(x) \underset{x\to+\infty}{\sim} \frac{1}{x^2}$.

3) On ne peut pas faire directement dans l'intégrale qui définit f le changement de variable $\theta = \tan t$ suggéré par l'énoncé, la fonction tan n'étant définie que sur des intervalles de la forme $\left]-\frac{\pi}{2}+k\pi\,;\frac{\pi}{2}+k\pi\right[$.

Cherchons déjà une primitive de $t \mapsto \dfrac{1}{x+\sin^2 t}$ sur tout intervalle de la forme $\left]-\frac{\pi}{2}+k\pi\,;\frac{\pi}{2}+k\pi\right[$, en posant $\theta = \tan t$.

$$\begin{aligned}\int \frac{\mathrm{d}t}{x+\sin^2 t} &= \int \frac{\mathrm{d}\theta}{(1+\theta^2)\left(x+1-\frac{1}{1+\theta^2}\right)} \\ &= \int \frac{\mathrm{d}\theta}{(x+1)\theta^2+x} \\ &= \frac{1}{\sqrt{x(x+1)}}\operatorname{Arctan}\left(\theta\sqrt{\frac{x+1}{x}}\right).\end{aligned}$$

Nous allons maintenant « découper » l'intégrale.

Soit N_x l'unique entier tel que

$$\frac{\pi}{2} + N_x\pi \leqslant \frac{1}{x} < \frac{\pi}{2} + (N_x+1)\pi\,.$$

Par la relation de Chasles,

$$\begin{aligned}f(x) = &\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{x+\sin^2 t} \\ &+ \sum_{k=1}^{N_x} \int_{-\frac{\pi}{2}+k\pi}^{\frac{\pi}{2}+k\pi} \frac{\mathrm{d}t}{x+\sin^2 t} \\ &+ \int_{\frac{\pi}{2}+N_x\pi}^{\frac{1}{x}} \frac{\mathrm{d}t}{x+\sin^2 t} \qquad (*)\end{aligned}$$

Or d'après le calcul fait auparavant :

$$\begin{aligned}&\int_{-\frac{\pi}{2}+k\pi}^{\frac{\pi}{2}+k\pi} \frac{\mathrm{d}t}{x+\sin^2 t} \\ &= \frac{1}{\sqrt{x(x+1)}}\left[\operatorname{Arctan}\left(\theta\sqrt{\frac{x+1}{x}}\right)\right]_{-\infty}^{+\infty} \\ &= \frac{\pi}{\sqrt{x(x+1)}}\,,\end{aligned}$$

et de même

$$\int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{x+\sin^2 t} = \frac{\frac{\pi}{2}}{\sqrt{x(x+1)}}$$

et enfin

$$0 \leqslant \int_{\frac{\pi}{2}+N_x\pi}^{\frac{1}{x}} \frac{\mathrm{d}t}{x+\sin^2 t} \leqslant \int_{\frac{\pi}{2}+N_x\pi}^{\frac{\pi}{2}+(N_x+1)\pi} \frac{\mathrm{d}t}{x+\sin^2 t} \leqslant \frac{\pi}{\sqrt{x(x+1)}}$$

donc finalement, en utilisant $(*)$, $f(x)$ est compris entre

$$\frac{1}{\sqrt{x(x+1)}}\left(\frac{\pi}{2}+N_x\pi\right)$$

et

$$\frac{1}{\sqrt{x(x+1)}}\left(\frac{\pi}{2}+(N_x+1)\pi\right).$$

Puisque

$$N_x = \left\lfloor \frac{\frac{1}{x}-\frac{\pi}{2}}{\pi} \right\rfloor \underset{x\to 0^+}{\sim} \frac{1}{\pi x},$$

on en déduit :

$$f(x) \underset{x\to 0^+}{\sim} \frac{1}{x\sqrt{x}}.$$

Exercice 10.7

• *Analyse*

Si f est une solution, on a :

$$f(x) = -1 - 2x\int_0^x f(t)\,\mathrm{d}t + \int_0^x tf(t)\,\mathrm{d}t$$

et cette écriture montre que f est de classe $\mathscr{C}^1$. Posons alors, pour tout $x\in\mathbb{R}$,

$$F(x) = \int_0^x f(t)\,\mathrm{d}t.$$

F est de classe $\mathscr{C}^2$ sur $\mathbb{R}$ et, en dérivant :

$$\begin{aligned} F''(x) &= f'(x) \\ &= -2\int_0^x f(t)\,\mathrm{d}t - 2xf(x) + xf(x) \\ &= -2F(x) - xF'(x) \end{aligned}$$

donc F est solution de l'équation différentielle :

$$y'' + xy' + 2y = 0,$$

avec les conditions initiales $F(0) = 0$ et $F'(0) = f(0) = -1$.

On cherche alors F comme somme d'une série entière de rayon de convergence $R > 0$, $F(x) = \sum_{n=0}^{+\infty} a_n x^n$. Puisque l'on peut dériver terme à terme une série entière sur son intervalle de convergence, on obtient, pour tout $x \in]-R\,;R[$,

$$\sum_{n=2}^{+\infty} n(n-1)a_n x^{n-2} + x\sum_{n=1}^{+\infty} na_n x^{n-1} + 2\sum_{n=0}^{+\infty} a_n x^n = 0$$

donc pour tout $n\in\mathbb{N}$:

$$(n+1)(n+2)a_{n+2} + (n+2)a_n = 0$$

soit : $a_{n+2} = -\dfrac{a_n}{n+1}$.

Puisque $a_0 = 0$ et $a_1 = -1$, on en déduit facilement $a_{2n} = 0$ pour tout entier n puis

$$\begin{aligned} a_{2n+1} &= \left(\frac{-1}{2n}\right)\left(\frac{-1}{2(n-1)}\right)\cdots\left(\frac{-1}{2}\right)a_1 \\ &= \frac{(-1)^{n+1}}{2^n n!}. \end{aligned}$$

Le rayon de convergence de cette série entière est clairement égal à $+\infty$, ce qui permet d'affirmer qu'elle définit bien une solution de l'équation différentielle précédente.

De plus, d'après le théorème de Cauchy-Lipschitz, il s'agit de la seule solution de cette équation différentielle qui vérifie les conditions initiales $y(0) = 0$ et $y'(0) = -1$, elle est donc égale à F.

Ainsi, pour tout $x\in\mathbb{R}$:

$$\begin{aligned} F(x) &= \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{2^n n!} x^{2n+1} \\ &= -x\sum_{n=0}^{+\infty} \frac{\left(-\frac{x^2}{2}\right)^n}{n!} = -x\mathrm{e}^{-\frac{x^2}{2}}. \end{aligned}$$

On en déduit :

$$\forall x\in\mathbb{R},\ f(x) = F'(x) = (x^2-1)\mathrm{e}^{-\frac{x^2}{2}}.$$

• *Synthèse*

Puisque nous avons procédé auparavant par conditions nécessaires, il reste à vérifier que la fonction f trouvée ci-dessus convient (il se pourrait que le problème n'ait pas de solution !), ce que nous vous laissons le soin de faire.

Exercice 10.8

1) Si $f(x_0) = 0$, puisque f est à valeurs positives, x_0 est un minimum (global) de f. On a donc $f'(x_0) = 0$.

La formule de Taylor-Young donne alors, au voisinage de x_0 :

$$f(x) = \frac{(x-x_0)^2}{2} f''(x_0) + o\big((x-x_0)^2\big)\,,$$

donc $f''(x_0) \geqslant 0$.

2) Par définition de m, on a

$$f(x_0) = f'(x_0) = \ldots = f^{(m-1)}(x_0) = 0,$$

donc la formule de Taylor avec reste intégrale donne :

$$\begin{aligned} &f(x_0+h)\\ &= \frac{1}{(m-1)!}\int_{x_0}^{x_0+h} (x_0+h-t)^{m-1} f^{(m)}(t)\,\mathrm{d}t\\ &= \frac{h^m}{(m-1)!}\int_0^1 (1-u)^{m-1} f^{(m)}(x_0+uh)\,\mathrm{d}u, \end{aligned}$$

en ayant fait le changement de variables $t = x_0 + hu$.
On pose alors :

$$g(h) = \tfrac{1}{(m-1)!}\int_0^1 (1-u)^{m-1} f^{(m)}(x_0+uh)\mathrm{d}u,$$

de sorte que $f(x_0+h) = h^m g(h)$, et le théorème de dérivation d'une intégrale à paramètre montre que g est de classe $\mathscr{C}^\infty$ (les hypothèses de domination s'obtiennent facilement en majorant les dérivées successives de f par leur borne supérieure sur des segments).

Par définition de m, on a $g(0) \neq 0$, et $f(x_0+h) \underset{h\to 0}{\sim} g(0)h^m$. f étant positive, on a forcément m pair et $g(0) > 0$.
Par suite, g reste > 0 au voisinage de 0, et dans ce voisinage on a :

$$\sqrt{f(x_0+h)} = \left|h^{\frac{m}{2}}\right| \sqrt{g(h)}.$$

Puisque g reste > 0 au voisinage de 0, $h \mapsto \sqrt{g(h)}$ est de classe $\mathscr{C}^\infty$ sur ce voisinage.

Compte tenu de la valeur absolue, $\sqrt{f}$ est localement de classe $\mathscr{C}^\infty$ si et seulement si $\frac{m}{2}$ est pair, c'est-à-dire m multiple de 4.

Exercice 10.9

• On applique l'inégalité de Taylor-Lagrange entre x et $x+h$:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} R_1\,,$$

puis entre x et $x-h$:

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2} R_2\,,$$

avec $|R_1| \leqslant M_2$ et $|R_2| \leqslant M_2$.
En faisant la différence de ces deux égalités, on obtient, pour $h \neq 0$:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h}{4}(R_1 - R_2),$$

d'où en utilisant l'inégalité triangulaire :

$$\forall\, h > 0,\ \big|f'(x)\big| \leqslant \frac{M_0}{h} + \frac{h}{2} M_2\,.$$

L'étude des variations sur $\mathbb{R}_+^*$ de la fonction $h \mapsto \frac{M_0}{h} + \frac{h}{2}M_2$ montre qu'elle admet un minimum égal à $\sqrt{2M_0M_2}$ (atteint pour $h = \sqrt{\frac{2M_0}{M_2}}$).
On a donc, pour tout $x \in \mathbb{R}$:

$$\big|f'(x)\big| \leqslant \sqrt{2M_0M_2}\,,$$

d'où le résultat.

Exercice 10.10

1) D'après les théorèmes usuels, φ est de classe $\mathscr{C}^n$ sur $I \setminus \{0\}$.
D'après la formule de Taylor-Young, compte tenu des hypothèses sur f, on a :

$$f(x) \underset{x\to 0}{=} o(x^n)\,.$$

– Donc $\varphi(x) = o\big(x^{n-1}\big)$, ce qui implique que $\lim\limits_{x\to 0} \varphi(x) = 0 = \varphi(0)$, donc φ est continue en 0, donc sur I.
– On a ensuite, pour $x \neq 0$, $x\varphi(x) = f(x)$ d'où :

$$x\varphi'(x) + \varphi(x) = f'(x) = o\big(x^{n-1}\big)\,,$$

puis $\varphi'(x) = o\big(x^{n-2}\big)$.
On en déduit que si $n \geqslant 2$, $\lim\limits_{x\to 0} \varphi'(x) = 0$, donc d'après le théorème de prolongement de la dérivée, φ est dérivable en 0 et $\varphi'(0) = \lim\limits_{x\to 0} \varphi'(x) = 0$, ce qui implique que φ' est continue en 0 donc sur I.
Finalement, φ est de classe $\mathscr{C}^1$.
– Supposons démontré jusqu'à un rang $p < n-1$ que φ est de classe $\mathscr{C}^p$ sur I et que $\varphi^{(p)}(x) = o\big(x^{n-p-1}\big)$.
Puisque $x\varphi(x) = f(x)$ pour $x \neq 0$, la formule de Leibniz donne, en dérivant $p+1$ fois le produit :

$$\begin{aligned} &x\varphi^{(p+1)}(x) + (p+1)\varphi^{(p)}(x)\\ &\qquad = f^{[p+1)}(x) = o\Big(x^{n-(p+1)}\Big) \end{aligned}$$

donc $x\varphi^{(p+1)}(x) = o\Big(x^{n-(p+1)}\Big)$ puis $\varphi^{(p+1)}(x) = o\Big(x^{n-p-2)}\Big)$
Puisque $n - p - 2 \geqslant 0$, on a $\lim\limits_{x\to 0} \varphi^{(p+1)}(x) = 0$, donc le théorème de prolongement de la dérivée appliqué à $\varphi^{(p)}$ montre que $\varphi^{(p)}$ est dérivable en 0, et que $\varphi^{(p+1)}(0) = \lim\limits_{x\to 0} \varphi^{(p+1)}(x) = 0$, c'est-à-dire que $\varphi^{(p+1)}$ est continue en 0 donc sur I et φ est de classe $\mathscr{C}^{p+1}$.

On a donc démontré par récurrence finie sur p que φ est de classe $\mathscr{C}^{p+1}$ pour $p < n-1$, donc de classe $\mathscr{C}^{n-1}$.

2) Soit g de classe $\mathscr{C}^n$ sur I. On considère alors la fonction f définie sur I par :

$$f(x) = g(x) - \left(g(0)+xg'(0)+\cdots+\frac{x^n}{n!}g^{(n)}(0)\right).$$

Alors f est de classe $\mathscr{C}^n$ sur I, et il est facile de vérifier que :

$$f(0) = f'(0) = \ldots = f^{(n)}(0) = 0\,.$$

En reprenant les résultats et les notations de la 1re question, on a alors , pour tout $x \in I$:

$$\psi(x) = \varphi(x) + \left(g'(0)+\cdots+\frac{x^{n-1}}{n!}g^{(n)}(0)\right),$$

donc ψ est aussi de classe $\mathscr{C}^{n-1}$ sur I.

Exercice 11.1

Puisque $\lim\limits_{n\to+\infty} \dfrac{(-1)^n}{n^\alpha} = 0$ (car $\alpha > 0$), on a :

$$\sin\left(\frac{(-1)^n}{n^\alpha}\right) \underset{n\to+\infty}{\sim} \frac{(-1)^n}{n^\alpha}\,,$$

donc

$$u_n \underset{n\to+\infty}{\sim} \frac{(-1)^n}{n^\alpha}\,,$$

mais cela ne permet pas de conclure car il ne s'agit pas de termes de signes constants !
Nous allons donc faire un développement limité de u_n afin de l'exprimer en fonctions de termes généraux de séries connues.

Puisque $\lim\limits_{n\to+\infty} \dfrac{(-1)^n}{n^\alpha} = 0$, on peut utiliser les développements limités usuels de $\sin x$ et de $\ln(1+x)$ et l'on obtient :

$$\begin{aligned} u_n &= \ln\left(1 + \frac{(-1)^n}{n^\alpha} + O\left(\frac{1}{n^{3\alpha}}\right)\right) \\ &= \frac{(-1)^n}{n^\alpha} - \frac{1}{2n^{2\alpha}} + O\left(\frac{1}{n^{3\alpha}}\right). \end{aligned}$$

On a donc trouvé $u_n = v_n + w_n$ avec :

$$v_n = \frac{(-1)^n}{n^\alpha} \text{ et } w_n = -\frac{1}{2n^{2\alpha}} + O\left(\frac{1}{n^{3\alpha}}\right).$$

La série de terme général v_n converge d'après le critère spécial sur les séries alternées.
On a aussi $w_n \underset{n\to+\infty}{\sim} -\dfrac{1}{2n^{2\alpha}}$, qui est le terme général d'une série *de signe constant*, et convergente si et seulement si $\alpha > \frac{1}{2}$. Il en est donc de même de la série de terme général w_n.

Puisque la somme de deux séries convergentes est convergente, et que celle d'une série convergente et d'une série divergente est divergente, on en déduit que :

la série de terme général

$$u_n = \ln\left(1 + \sin\left(\frac{(-1)^n}{n^\alpha}\right)\right)$$

converge si et seulement si $\alpha > \dfrac{1}{2}$.

Exercice 11.2

On commence par effectuer un développement limité du dénominateur. Il se pose alors le problème habituel : à quel ordre effectuer le développement limité ? L'idée de base est simple : il faut que, dans l'expression *finale* de u_n, on puisse conclure quant à la nature des différentes séries qui apparaissent. Il faut donc (en général) que le dernier terme du développement limité final soit en $o\left(\dfrac{1}{n^\alpha}\right)$ ou en $O\left(\dfrac{1}{n^\alpha}\right)$ avec $\alpha > 1$.
Pour l'exemple qui suit, on prendra $\alpha = 2$. Mais les développement limités intermédiaires seront faits à un ordre supérieur, puisque l'on va être amené à multiplier et diviser par des puissances de n.

$$\begin{aligned}
&\ln(n^2+n+1)-\ln(n^2+1)\\
&\quad=\ln\left(\frac{n^2+n+1}{n^2+1}\right)=\ln\left(1+\frac{n}{n^2+1}\right)\\
&\quad=\ln\left(1+\frac{n}{n^2\left(1+\frac{1}{n^2}\right)}\right)\\
&\quad=\ln\left(1+\frac{1}{n\left(1+\frac{1}{n^2}\right)}\right),
\end{aligned}$$

puis en utilisant le développement limité $\frac{1}{1+x}\underset{x\to0}{=}1-x+O\big(x^2\big)$:

$$\begin{aligned}
&\ln(n^2+n+1)-\ln(n^2+1)\\
&\qquad=\ln\left(1+\frac{1}{n}\left(1-\frac{1}{n^2}+O\left(\frac{1}{n^4}\right)\right)\right)\\
&\qquad=\ln\left(1+\frac{1}{n}-\frac{1}{n^3}+O\left(\frac{1}{n^5}\right)\right)\\
&\qquad=\ln\left(1+\frac{1}{n}-\frac{1}{n^3}+O\left(\frac{1}{n^4}\right)\right),
\end{aligned}$$

et enfin en utilisant le développement limité du ln à l'ordre 4, soit :

$$\ln(1+x)\underset{x\to0}{=}x-\frac{x^2}{2}+\frac{x^3}{3}+O\big(x^4\big)\,,$$

on obtient finalement :

$$\begin{aligned}
&\ln(n^2+n+1)-\ln(n^2+1)\\
&\qquad=\left(\frac{1}{n}-\frac{1}{n^3}\right)-\frac{1}{2}\left(\frac{1}{n}-\frac{1}{n^3}\right)^2\\
&\qquad\quad+\frac{1}{3}\left(\frac{1}{n}-\frac{1}{n^3}\right)^3+O\left(\frac{1}{n^4}\right)\\
&\qquad=\frac{1}{n}-\frac{1}{n^3}-\frac{1}{2n^2}+\frac{1}{3n^3}+O\left(\frac{1}{n^4}\right)\\
&\qquad=\frac{1}{n}-\frac{1}{2n^2}-\frac{2}{3n^3}+O\left(\frac{1}{n^4}\right).
\end{aligned}$$

On en déduit :

$$\begin{aligned}
&\frac{1}{\ln(n^2+n+1)-\ln(n^2+1)}\\
&\qquad=\frac{1}{\frac{1}{n}-\frac{1}{2n^2}-\frac{2}{3n^3}+O\left(\frac{1}{n^4}\right)}\\
&\qquad=\frac{n}{1-\left(\frac{1}{2n}+\frac{2}{3n^2}+O\left(\frac{1}{n^3}\right)\right)}
\end{aligned}$$

puis on utilise le développement limité de $\frac{1}{1-x}$, à l'ordre 3 puisque le développement limité précédent est aussi à l'ordre 3 :

$$\frac{1}{1-x}\underset{x\to0}{=}1+x+x^2+O\big(x^3\big)$$

d'où l'on tire :

$$\begin{aligned}
&\frac{1}{\ln(n^2+n+1)-\ln(n^2+1)}\\
&\quad=n\left(1+\left(\frac{1}{2n}+\frac{2}{3n^2}\right)+\left(\frac{1}{2n}+\frac{2}{3n^2}\right)^2\right.\\
&\qquad\qquad\left.+O\left(\frac{1}{n^3}\right)\right)\\
&\quad=n\left(1+\frac{1}{2n}+\frac{2}{3n^2}+\frac{1}{4n^2}+O\left(\frac{1}{n^3}\right)\right)\\
&\quad=n+\frac{1}{2}+\frac{11}{12n}+O\left(\frac{1}{n^2}\right)
\end{aligned}$$

et on obtient enfin :

$$u_n=-\frac{11}{12n}+O\left(\frac{1}{n^2}\right).$$

Il ne reste plus qu'à conclure :

• $u_n\underset{n\to+\infty}{\sim}-\dfrac{11}{12n}$; puisque la série de terme général $\dfrac{1}{n}$ diverge, il en est de même de la série $\sum u_n$.

• $(-1)^n u_n=-\dfrac{11}{12}\cdot\dfrac{(-1)^n}{n}+O\left(\dfrac{1}{n^2}\right)$.

Puisque la série de terme général $\dfrac{(-1)^n}{n}$ converge (série harmonique alternée, qui vérifie le critère spécial sur les séries alternées) et qu'une série dont le terme général est un $O\left(\dfrac{1}{n^2}\right)$ est absolument convergente, la série de terme général $(-1)^n u_n$ converge, comme somme de deux séries convergentes.

En conclusion, la série de terme général u_n diverge et celle de terme général $(-1)^n u_n$ converge.

✓ Même si un simple équivalent suffisait pour conclure quant à la nature de $\sum u_n$, il était indispensable d'avoir un développement limité plus complet pour l'étude de $\sum(-1)^n u_n$.

✓ Vous trouverez peut-être le calcul précédent un peu long. Nous avons volontairement détaillé toutes les étapes, afin que vous compreniez les mécanismes à mettre en œuvre lors d'un développement limité.

Exercice 11.3

On montre déjà par récurrence que $u_n>0$ pour tout n, ce qui permet de vérifier en

même temps que la définition a bien un sens.
Cela implique que la suite u est croissante.

• Si la suite (u_n) converge alors elle est bornée. Soit M un majorant (positif) des termes de la suite u. On a pour tout n :

$$0 \leqslant a_n = u_n(u_{n+1} - u_n) \leqslant M(u_{n+1} - u_n)$$

et puisque la série de terme général $u_{n+1} - u_n$ converge, il en est de même de $\sum a_n$ par comparaison de séries à termes positifs.

• Si $\sum a_n$ converge, alors

$$0 \leqslant u_{n+1} - u_n = \frac{a_n}{u_n} \leqslant \frac{a_n}{u_0}$$

donc $\sum(u_{n+1} - u_n)$ converge par comparaison de séries à termes positifs, d'où la suite $(u_n)_{n\in\mathbb{N}}$ converge.

Exercice 11.4

1) Pour $n \geqslant 1$,

$$\begin{aligned}\ln s_n - \ln s_{n-1} &= -\ln\left(\frac{s_{n-1}}{s_n}\right)\\ &= -\ln\left(\frac{s_n - u_n}{s_n}\right)\\ &= -\ln\left(1 - \frac{u_n}{s_n}\right).\end{aligned}$$

On distingue alors deux cas.

– Si $\lim\limits_{n\to+\infty} \dfrac{u_n}{s_n} = 0$ alors :

$$-\ln\left(1 - \frac{u_n}{s_n}\right) \underset{n\to+\infty}{\sim} \frac{u_n}{s_n}$$

donc $\dfrac{u_n}{s_n} \underset{n\to+\infty}{\sim} \ln s_n - \ln s_{n-1}$.

Or la série de terme général $\ln s_n - \ln s_{n-1}$ diverge, car

$$\sum_{n=1}^{N}(\ln s_n - \ln s_{n-1}) = \ln s_N - \ln s_0$$

tend vers $+\infty$ avec N , puisque $\sum u_n$ est une série à termes positifs divergente.

Il en résulte que la série $\sum \dfrac{u_n}{s_n}$ diverge, par comparaison de séries à termes positifs.

– Si $\dfrac{u_n}{s_n}$ ne tend pas vers zéro, la série $\sum \dfrac{u_n}{s_n}$ diverge grossièrement !

2) On étudie 3 cas.

– Si $\alpha \leqslant 0$.

$\dfrac{u_n}{s_n^\alpha} \geqslant \dfrac{u_n}{s_0^\alpha}$ (puisque la suite (s_n) est croissante), donc la série $\sum \dfrac{u_n}{s_n^\alpha}$ diverge d'après le théorème de comparaison de séries à termes positifs.

– Si $0 < \alpha \leqslant 1$.

Puisque $\lim\limits_{n\to+\infty} s_n = +\infty$, il existe un rang à partir duquel on aura $s_n \geqslant 1$ et alors $s_n^\alpha \leqslant s_n$ donc $\dfrac{u_n}{s_n^\alpha} \geqslant \dfrac{u_n}{s_n}$ et d'après la question **1)**, la série $\sum \dfrac{u_n}{s_n^\alpha}$ diverge.

– Si $\alpha > 1$.

La fonction $t \mapsto \dfrac{1}{t^\alpha}$ étant décroissante, on a :

$$\frac{u_n}{s_n^\alpha} = \frac{s_n - s_{n-1}}{s_n^\alpha} \leqslant \int_{s_{n-1}}^{s_n} \frac{\mathrm{d}t}{t^\alpha},$$

donc, en calculant cette intégrale :

$$\frac{u_n}{s_n^\alpha} \leqslant \frac{1}{\alpha - 1}\left(\frac{1}{s_{n-1}^{\alpha-1}} - \frac{1}{s_n^{\alpha-1}}\right).$$

Or la série télescopique de terme général $\dfrac{1}{s_{n-1}^{\alpha-1}} - \dfrac{1}{s_n^{\alpha-1}}$ est convergente puisque $\lim\limits_{n\to+\infty} \dfrac{1}{s_n^{\alpha-1}} = 0$. Par comparaison, il en est de même de la série $\sum \dfrac{u_n}{s_n^\alpha}$.

Exercice 11.5

1) La fonction $x \mapsto \dfrac{\ln x}{x}$ est décroissante pour $x \geqslant \mathrm{e}$ (calculer sa dérivée) donc la suite $n \mapsto f(n)$ est décroissante à partir du rang 3. Puisque de plus $\lim\limits_{n\to+\infty} \dfrac{\ln n}{n} = 0$, la série de terme général $(-1)^n \dfrac{\ln n}{n}$ vérifie le critère spécial sur les séries alternées, donc converge.

2) • Pour $n \geqslant 2$, posons :

$$w_n = \int_{n-1}^{n} f(t)\,\mathrm{d}t - f(n).$$

On sait d'après le cours (**[S11.7]**) que la série $\sum w_n$ converge.

• On a :

$$\begin{aligned}
&2\sum_{k=1}^{n} f(2k) - \sum_{k=1}^{2n} f(k) \\
&\quad = 2\sum_{k=1}^{n} f(2k) - \sum_{k=1}^{n} f(2k) - \sum_{k=1}^{n} f(2k-1) \\
&\quad = \sum_{k=1}^{n} f(2k) - \sum_{k=1}^{n} f(2k-1) \\
&\quad = \sum_{k=1}^{2n} (-1)^k f(k)\,.
\end{aligned}$$

• Puisque $2f(2k) = \dfrac{\ln 2}{k} + f(k)$, on en tire :

$$\sum_{k=1}^{2n} (-1)^k f(k) = \ln 2 \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=n+1}^{2n} f(k) \quad (*)$$

Or :

$$\begin{aligned}
&\sum_{k=n+1}^{2n} f(k) = \sum_{k=n+1}^{2n} \left(\int_{k-1}^{k} f(t)\,\mathrm{d}t - w_k \right) \\
&= \int_{n}^{2n} f(t)\,\mathrm{d}t - \sum_{k=n+1}^{2n} w_k \\
&= \left[\frac{(\ln t)^2}{2} \right]_n^{2n} - \sum_{k=n+1}^{2n} w_k \\
&= \frac{(\ln(2n) - \ln n)(\ln(2n) + \ln n)}{2} - \sum_{k=n+1}^{2n} w_k \\
&= \frac{(\ln 2)^2}{2} + \ln 2 \ln n - \sum_{k=n+1}^{2n} w_k\,.
\end{aligned}$$

D'autre part, on sait que :

$$\sum_{k=1}^{n} \frac{1}{k} = \ln n + \gamma + o(1)\ ,$$

où γ est la constante d'Euler, donc en remplaçant toutes les expressions précédentes dans $(*)$:

$$\begin{aligned}
\sum_{k=1}^{2n} (-1)^k f(k) &= \gamma \ln 2 - \frac{(\ln 2)^2}{2} \\
&\quad + \sum_{k=n+1}^{2n} w_k + o(1)\ .
\end{aligned}$$

Mais la série de terme général w_n étant convergente, on a $\lim\limits_{n\to+\infty} \sum\limits_{k=n+1}^{2n} w_k = 0$, donc finalement :

$$\sum_{k=1}^{+\infty} (-1)^k \frac{\ln k}{k} = \gamma \ln 2 - \frac{\ln^2 2}{2}\,.$$

Exercice 11.6

La série proposée est celle de terme général :

$$\begin{aligned}
u_n &= \sum_{k=0}^{n} \frac{(-1)^k}{(k+1)^\alpha} \cdot \frac{(-1)^{n-k}}{(n-k+1)^\beta} \\
&= (-1)^n \sum_{k=0}^{n} \frac{1}{(k+1)^\alpha (n-k+1)^\beta}\,.
\end{aligned}$$

Pour $x \in [0\,;n]$ posons :

$$f(x) = (x+1)^\alpha (n-x+1)^\beta.$$

On calcule :

$$\begin{aligned}
f'(x) &= \alpha(x+1)^{\alpha-1}(n-x+1)^\beta \\
&\qquad - \beta(x+1)^\alpha (n-x+1)^{\beta-1} \\
&= (x+1)^{\alpha-1}(n-x+1)^{\beta-1} \\
&\qquad \times (\alpha n + \alpha - \beta - (\alpha+\beta)x)\,,
\end{aligned}$$

ce qui montre que f est d'abord croissante puis décroissante, son maximum étant atteint en $x = \dfrac{\alpha n + \alpha - \beta}{\alpha + \beta}$. Un petit calcul montre ensuite que ce maximum est égal à $\dfrac{\alpha^\alpha \beta^\beta}{\alpha+\beta}(n+2)^{\alpha+\beta}$.

On aura donc, pour tout $k \in [\![0\,;n]\!]$:

$$(k+1)^\alpha (n-k+1)^\beta \leqslant \frac{\alpha^\alpha \beta^\beta}{\alpha+\beta}(n+2)^{\alpha+\beta}$$

d'où : $|u_n| \geqslant \dfrac{n+1}{(n+2)^{\alpha+\beta}} \dfrac{\alpha+\beta}{\alpha^\alpha \beta^\beta}$,

et finalement, lorsque $\alpha+\beta \leqslant 1$, cette quantité ne tend pas vers 0 lorsque $n \to +\infty$, ce qui montre que la série de terme général u_n diverge grossièrement.

Exercice 11.7

Pour $n \in \mathbb{N}^*$ on a :

$$\ln u_n = \ln(n!) + n - \left(n + \frac{1}{2}\right) \ln n\,,$$

donc, près un calcul simple :

$$\begin{aligned}
&\ln\left(\frac{u_n}{u_{n+1}}\right) = \left(n+\frac{1}{2}\right) \ln\left(\frac{n+1}{n}\right) - 1 \\
&= -1 + \left(n + \tfrac{1}{2}\right) \\
&\qquad \times \left(\tfrac{1}{n} - \tfrac{1}{2n^2} + \tfrac{1}{3n^3} - \tfrac{1}{4n^4} + O\left(\tfrac{1}{n^5}\right)\right) \\
&= \frac{1}{12n^2} - \frac{1}{12n^3} + O\left(\frac{1}{n^4}\right).
\end{aligned}$$

D'autre part, un autre développement limité donne :

$$\frac{1}{n(n+1)} = \frac{1}{n^2}\left(1+\frac{1}{n}\right)^{-1} = \frac{1}{n^2} - \frac{1}{n^3} + O\left(\frac{1}{n^4}\right),$$

de sorte que :

$$\ln\left(\frac{u_n}{u_{n+1}}\right) = \frac{1}{12n(n+1)} + O\left(\frac{1}{n^4}\right).$$

Puisque la série de terme général $\frac{1}{12n(n+1)}$ est convergente, cela prouve déjà que la série de terme général $\ln\left(\frac{u_n}{u_{n+1}}\right)$ l'est aussi, et puisque $\sum \frac{1}{n^4}$ est une série à termes positifs convergente, le théorème de sommation pour les restes de séries convergentes donne :

$$\sum_{k=n}^{+\infty} \ln\left(\frac{u_k}{u_{k+1}}\right) - \frac{1}{12k(k+1)} = O\left(\sum_{k=n}^{+\infty} \frac{1}{k^4}\right).$$

Par une classique méthode de comparaison série-intégrale (par exemple, ou voir aussi **[SF11.20]** page 470), on a :

$$\sum_{k=n}^{+\infty} \frac{1}{k^4} \underset{n\to+\infty}{\sim} \int_n^{+\infty} \frac{\mathrm{d}t}{t^4} = \frac{1}{3n^3}.$$

D'autre part, par télescopage :

$$\sum_{k=n}^{+\infty} \frac{1}{k(k+1)} = \sum_{k=n}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{n}.$$

Enfin, par la formule de Stirling, on a $\lim\limits_{n\to+\infty} u_n = \sqrt{2\pi}$ donc :

$$\sum_{k=n}^{+\infty} \ln\left(\frac{u_k}{u_{k+1}}\right) = \ln u_n - \ln(\sqrt{2\pi}).$$

En rassemblant tous ces résultats, on obtient :

$$\ln u_n = \ln(\sqrt{2\pi}) + \frac{1}{12n} + O\left(\frac{1}{n^3}\right),$$

soit :

$$u_n = \sqrt{2\pi}\mathrm{e}^{\frac{1}{12n}+O\left(\frac{1}{n^3}\right)} = \sqrt{2\pi}\left(1+\frac{1}{12n}+\frac{1}{2.12^2.n^2}+O\left(\frac{1}{n^3}\right)\right).$$

Puisque $n! = u_n\left(\frac{n}{\mathrm{e}}\right)^n\sqrt{n}$ on en déduit :

$$n!\left(\frac{n}{\mathrm{e}}\right)^n\sqrt{2\pi n} \times \left(1+\frac{1}{12n}+\frac{1}{288n^2}+O\left(\frac{1}{n^3}\right)\right).$$

Exercice 11.8

Notons pour tout $(p,q) \in \mathbb{N}^2$:

$$u_{p,q} = \frac{1}{p!q!(p+q+1)}.$$

On peut alors considérer la partition de $\mathbb{N}^2$ formée par les ensembles :

$$I_n = \left\{(p,q) \in \mathbb{N}^2 \mid p+q=n\right\}.$$

Pour tout $n \in \mathbb{N}$ on a :

$$\sum_{(p,q)\in I_n} u_{p,q} = \frac{1}{n+1}\sum_{p+q=n}\frac{1}{p!q!} = \frac{1}{n+1}\sum_{p+q=n}\frac{1}{(p+q)!}\binom{p+q}{q} = \frac{1}{(n+1)!}\sum_{p+q=n}\binom{p+q}{q} = \frac{1}{(n+1)!}2^n,$$

à l'aide de la formule du binôme.

Puisque la série de terme général $\frac{1}{(n+1)!}2^n$ converge, la famille $(u_{p,q})$, à termes positifs, est sommable d'après le théorème de sommation par paquets. De plus :

$$\sum_{(p,q)\in\mathbb{N}^2} u_{p,q} = \sum_{n=0}^{+\infty}\frac{2^n}{(n+1)!} = \frac{1}{2}\sum_{n=0}^{+\infty}\frac{2^{n+1}}{(n+1)!} = \frac{1}{2}\sum_{n=1}^{+\infty}\frac{2^n}{n!} = \frac{1}{2}(\mathrm{e}^2-1).$$

Exercice 11.9

Pour tout entier $n \geqslant 2$ et tout complexe z on a :

$$\left(\zeta(q)-1\right)z^q = \sum_{p=2}^{+\infty}\frac{z^q}{p^q}.$$

On considère donc la suite double $(u_{p,q})$ avec $u_{p,q} = \frac{z^q}{p^q}$ pour $p,\, q \geqslant 2$.

Pour $p \geqslant 2$ fixé, la série $\sum\limits_{q\geqslant 2}|u_{p,q}|$ converge, puisque c'est une série géométrique de raison $|z|\,p$ et que l'on a supposé $|z|<2$.

De plus :

$$v_p = \sum_{q=2}^{+\infty} \left|\frac{z}{p}\right|^q = \left|\frac{z}{p}\right|^2 \frac{1}{1-\left|\frac{z}{p}\right|} = \frac{|z|^2}{p(p-|z|)}\,.$$

Puisque $v_p \underset{p\to+\infty}{\sim} \dfrac{|z|^2}{p^2}$, la série de terme général v_p est convergente.
Il résulte alors du théorème de Fubini que la suite double $(u_{p,q})_{p,q\geqslant 2}$ est sommable.
Ce même théorème permet aussi d'écrire, si l'on note $S = \sum\limits_{p,q\geqslant 2} u_{p,q}$:

$$S = \sum_{q=2}^{+\infty}\left(\sum_{p=2}^{+\infty}\frac{z^q}{p^q}\right) = \sum_{q=2}^{+\infty}\big(\zeta(q)-1\big)\,z^q$$

et

$$S = \sum_{p=2}^{+\infty}\left(\sum_{q=2}^{+\infty}\frac{z^q}{p^q}\right) = \sum_{p=2}^{+\infty}\frac{z^2}{p(p-z)}\,,$$

d'où l'égalité cherchée.

Exercice 11.10

1) Le théorème de Fubini s'applique ici sans problème, et donne :

$$\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{p^2q^2} = \left(\sum_{n=1}^{+\infty}\frac{1}{n^2}\right)^2 = \frac{\pi^4}{36}\,.$$

2) • Notons déjà que la famille proposée est sommable, puisqu'elle est à termes positifs et que ses « sommes partielles » sont majorées par $\dfrac{\pi^4}{36}$.

• On considère alors la partition de $(\mathbb{N}^*)^2$ formée des ensembles :

$$I_n = \big\{(p,q)\in(\mathbb{N}^*)^2 \;\big|\; p\wedge q = n\big\}$$

pour $n\in\mathbb{N}^*$. Puisque la famille est sommable, le théorème de sommation par paquets donne :

$$\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{p^2q^2} = \sum_{n=1}^{+\infty}\left(\sum_{p\wedge q=n}\frac{1}{p^2q^2}\right).$$

Mais $p\wedge q = n$ si et seulement si il existe p' et q' entiers tels que :

$$p = np'\,,\ q = nq' \quad\text{et } p'\wedge q' = 1\,,$$

donc :

$$\sum_{p\wedge q=n}\frac{1}{p^2q^2} = \sum_{p'\wedge q'=1}\frac{1}{(np')^2(nq')^2} = \frac{1}{n^4}\sum_{p\wedge q=1}\frac{1}{p^2q^2}\,.$$

Finalement on a :

$$\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{p^2q^2} = \left(\sum_{n=1}^{+\infty}\frac{1}{n^4}\right)\left(\sum_{p\wedge q=1}\frac{1}{p^2q^2}\right),$$

d'où :

$$\sum_{p\wedge q=1}\frac{1}{p^2q^2} = \frac{\frac{\pi^4}{36}}{\frac{\pi^4}{90}} = \frac{90}{36} = \frac{5}{2}\,.$$

Exercice 11.11

Pour $|z|<1$ on a $\dfrac{z^n}{1-z^n} = \sum\limits_{k=1}^{+\infty} z^{nk}$.

On considère donc la suite double $(u_{n,k})$ définie par :

$$\forall\,(n,k)\in(\mathbb{N}^*)^2,\ u_{n,k} = z^{nk}\,.$$

Pour tout $n\in\mathbb{N}^*$, la série $\sum\limits_{k\geqslant 1}|u_{n,k}|$ est convergente et :

$$v_n = \sum_{k=1}^{+\infty}|u_{n,k}| = \sum_{k=1}^{+\infty}|z|^{nk} = \frac{|z|^n}{1-|z|^n}\,.$$

Puisque $|z|<1$, on a $\dfrac{|z|^n}{1-|z|^n} \underset{n\to+\infty}{\sim} |z|^n$, donc la série $\sum v_n$ est convergente, par comparaison à une série géométrique.
Il résulte alors du théorème de Fubini que la famille $(u_{n,k})_{(n,k)\in(\mathbb{N}^*)^2}$ est sommable, et :

$$\sum_{(n,k)\in(\mathbb{N}^*)^2}u_{n,k} = \sum_{n=1}^{+\infty}\frac{z^n}{1-z^n}\,.$$

On peut alors utiliser le théorème de sommation par paquets. En réalisant une partition de $(\mathbb{N}^*)^2$ selon les ensembles :

$$I_p = \big\{(n,k)\in(\mathbb{N}^*)^2 \;\big|\; nk = p\big\} \quad (p\in\mathbb{N}^*),$$

on obtient :

$$\sum_{(n,k)\in(\mathbb{N}^*)^2} z^{nk} = \sum_{p=1}^{+\infty}\left(\sum_{nk=p} z^{nk}\right) = \sum_{p=1}^{+\infty}\mathrm{Card}(I_p)z^p\,.$$

Mais le cardinal de I_p, c'est-à-dire le nombre de couples (n,k) d'entiers non nuls tels que $nk=p$ est exactement égal au nombre $d(p)$ de diviseurs de p, donc finalement :

$$\sum_{n=1}^{+\infty}\frac{z^n}{1-z^n} = \sum_{p=1}^{+\infty}d(p)z^p\,.$$

Exercice 11.12

• Pour tout entier n, $f\colon t \mapsto \dfrac{\ln t}{1+t^n}$ est continue sur $]0\,;1]$, et est majorée en valeur absolue par $|\ln t|$ qui est intégrable sur $]0\,;1]$; cela assure l'existence de la suite proposée.

• Fixons $n \in \mathbb{N}$, $n \geqslant 2$. Pour tout $t \in [0\,;1[$, on sait que $\dfrac{1}{1+t^n} = \sum_{p=0}^{+\infty} (-1)^p t^{np}$.

Cela suggère d'introduire la suite de fonctions $(f_p)_{p\in\mathbb{N}}$ définies par :

$$\forall\, t \in]0\,;1],\ f_p(t) = (-1)^p t^{np} \ln t\,.$$

La série $\sum_{p\in\mathbb{N}} f_p$ converge simplement vers la fonction f sur $]0\,;1[$. En outre, chaque f_p est continue et intégrable sur $]0\,;1[$.
Un calcul simple (intégration par parties) donne $\int_0^1 |f_p(t)|\,\mathrm{d}t = \dfrac{1}{(1+np)^2}$. La série $\sum_{p\in\mathbb{N}} \int_0^1 |f_p|$ est donc convergente, et le théorème d'intégration terme à terme assure alors que :

$$u_n = \sum_{p=0}^{+\infty} \int_0^1 f_p(t)\,\mathrm{d}t = \sum_{p=0}^{+\infty} \frac{(-1)^{p+1}}{(1+np)^2}\,.$$

• On connaît le développement en série entière sur $]-1\,;1[$:

$$\frac{1}{(1+x)^2} = \sum_{k=1}^{+\infty} (-1)^{k-1} k x^{k-1}\,,$$

obtenu par exemple par dérivation de celui de $\dfrac{1}{1+x}$. On en déduit :

$$\begin{aligned} u_n &= -1 + \sum_{p=1}^{+\infty} \frac{(-1)^{p+1}}{(np)^2 \left(1+\frac{1}{np}\right)^2} \\ &= -1 + \sum_{p=1}^{+\infty} \sum_{k=1}^{+\infty} \frac{(-1)^{k+p}\,k}{(np)^{k+1}} \end{aligned}$$

(en effet, puisque l'on a supposé $n \geqslant 2$, on a bien $\dfrac{1}{np} < 1$, ce qui permet d'utiliser le développement en série entière précédent).
Cela justifie l'introduction de la suite double donnée par :

$$a_{p,k} = \frac{(-1)^{k+p} k}{(np)^{k+1}}\,,$$

pour $(p,k) \in (\mathbb{N}^*)^2$.

On a alors, pour tout $p \geqslant 1$, en utilisant des calculs semblables aux précédents :

$$\sum_{k=1}^{+\infty} |a_{p,k}| = \frac{1}{(1-np)^2} \underset{p\to+\infty}{\sim} \frac{1}{n^2 p^2}\,.$$

Ainsi, à p fixé la série $\sum_{k\geqslant 1} |a_{p,k}|$ converge et la série $\sum_{p\geqslant 1} \left(\sum_{k=1}^{+\infty} |a_{p,k}| \right)$ converge.
D'après le théorème de Fubini, la suite double $(a_{p,k})_{(p,k)\in(\mathbb{N}^*)^2}$ est sommable et par interversion des sommations, on peut écrire, pour tout $n \geqslant 2$:

$$(*) \qquad u_n = -1 + \sum_{k=2}^{+\infty} \frac{a_k}{n^k}\,,$$

où l'on a posé :

$$\forall\, k \geqslant 2,\ a_k = (-1)^k (k-1) \sum_{p=1}^{+\infty} \frac{(-1)^{p-1}}{p^k}\,.$$

La relation $(*)$ fournit par troncature de la série un développement limité à tout ordre selon les puissances de $\dfrac{1}{n}$ de u_n.

Exercice 12.1

En appliquant la définition de la convergence uniforme avec $\varepsilon = 1$, on obtient qu'il existe $N \in \mathbb{N}$ tel que :

$$\forall\, n \geqslant N,\ \|P_n - f\|_\infty \leqslant 1\,.$$

En utilisant l'inégalité triangulaire on aura donc, pour tout $n \geqslant N$:

$$\|P_n - P_N\|_\infty \leqslant 2\,.$$

$P_n - P_N$ étant une fonction polynôme bornée sur $\mathbb{R}$, elle est constante :

$$\exists\, \lambda_n \in \mathbb{R},\ P_n = P_N + \lambda_n\,.$$

En particulier, $\lambda_n = P_n(0) - P_N(0)$ donc $\lim\limits_{n\to+\infty} \lambda_n = f(0) - P_N(0)$ existe.
Notons λ cette limite. On a alors, pour tout x réel :

$$\begin{aligned} f(x) &= \lim_{n\to+\infty} P_n(x) = \lim_{n\to+\infty} \big(P_N(x) + \lambda_n\big) \\ &= P_N(x) + \lambda\,, \end{aligned}$$

donc $f = P_N + \lambda$ est une fonction polynôme.

Exercice 12.2

1) a) Pour tout $x \in \mathbb{R}$ et tout $n \in \mathbb{N}$, on a, grâce à l'inégalité triangulaire :

$$\begin{aligned}&\big|f_n\big(g_n(x)\big) - f\big(g(x)\big)\big| \\ &\quad\leqslant \big|f_n\big(g_n(x)\big) - f\big(g_n(x)\big)\big| \\ &\qquad\qquad + \big|f\big(g_n(x)\big) - f\big(g(x)\big)\big| \\ &\quad\leqslant \|f_n - f\|_\infty + \big|f\big(g_n(x)\big) - f\big(g(x)\big)\big|\,.\end{aligned}$$

Or :

$$\lim_{n\to+\infty} \|f_n - f\|_\infty = 0$$

puisque (f_n) converge uniformément vers f, et :

$$\lim_{n\to+\infty} f\big(g_n(x)\big) - f\big(g(x)\big) = 0$$

puisque f est continue au point $g(x)$ (comme limite uniforme d'une suite de fonctions continues).
On en déduit :

$$\lim_{n\to+\infty} f_n\big(g_n(x)\big) - f\big(g(x)\big) = 0\,,$$

c'est-à-dire que la suite $(f_n \circ g_n)$ converge simplement vers $f \circ g$.

b) Considérons la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ définies sur $\mathbb{R}$ par :

$$\forall\, x \leqslant 0,\ f_n(x) = 0\ ;\ \forall\, x \geqslant \frac{2}{n},\ f_n(x) = 0;$$
$$f_n\left(\frac{1}{n}\right) = n\ ;$$

f_n continue affine par morceaux.

La suite (f_n) converge simplement sur $\mathbb{R}$ vers la fonction nulle ; en effet, cela est vrai pour $x \leqslant 0$, et si $x > 0$, il existe $N \in \mathbb{N}^*$ tel que $x \geqslant \dfrac{2}{N}$ et on a alors pour $n \geqslant N$, $f_n(x) = 0$.
Si l'on considère maintenant la suite $(g_n)_{n\in\mathbb{N}^*}$ de fonctions constantes telles que $g_n(x) = \dfrac{1}{n}$ pour tout x, la suite $(g_n)_{n\in\mathbb{N}^*}$ converge (uniformément) vers la fonction nulle, mais la suite $(f_n \circ g_n)$ n'a pas de limite simple puisque $f_n \circ g_n(x) = n$ pour tout x.

2) a) Soit $\varepsilon > 0$ fixé. L'uniforme continuité de f assure l'existence d'un $\alpha > 0$ tel que pour tout $(x,y) \in \mathbb{R}^2$:

$$|x - y| \leqslant \alpha \implies |f(x) - f(y)| \leqslant \frac{\varepsilon}{2}\cdot$$

La convergence uniforme de (g_n) vers g s'écrit :

$$\begin{aligned}&\exists\, N \in \mathbb{N},\ n \geqslant N \implies \\ &\qquad\qquad \forall\, x \in \mathbb{R},\ |g_n(x) - g(x)| \leqslant \alpha,\end{aligned}$$

donc pour $n \geqslant N$ on a :

$$\forall\, x \in \mathbb{R},\ \big|f\big(g_n(x)\big) - f\big(g(x)\big)\big| \leqslant \frac{\varepsilon}{2}$$

Enfin, la convergence uniforme de (f_n) vers f s'écrit :

$$\exists\, N' \in \mathbb{N},\ n \geqslant N' \implies \|f_n - f\|_\infty \leqslant \frac{\varepsilon}{2}\,,$$

donc pour $n \geqslant \max(N,N')$ on a, pour tout $x \in \mathbb{R}$:

$$\begin{aligned}&\big|f_n\big(g_n(x)\big) - f\big(g(x)\big)\big| \\ &\quad\leqslant \|f_n - f\|_\infty + \big|f\big(g_n(x)\big) - f\big(g(x)\big)\big| \\ &\quad\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,\end{aligned}$$

ce qui signifie que la suite $(f_n \circ g_n)$ converge uniformément vers $f \circ g$ sur $\mathbb{R}$.

b) Considérons la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ définies sur $\mathbb{R}$ par :

$$\forall\, x \in \mathbb{R},\ f_n(x) = x^2 + \frac{1}{n}\cdot$$

La suite $(f_n)_{n\in\mathbb{N}^*}$ converge uniformément sur $\mathbb{R}$ vers $f : x \mapsto x^2$.
D'après la question **1)**, la suite $(f_n \circ f_n)$ converge simplement sur $\mathbb{R}$ vers la fonction $f \circ f : x \mapsto x^4$.
Cependant, la convergence n'est pas uniforme sur $\mathbb{R}$. En effet, si l'on pose $h_n = f_n \circ f_n - f \circ f$ on a, pour tout $x \in \mathbb{R}$:

$$\begin{aligned}h_n(x) &= \left(x^2 + \frac{1}{n}\right)^2 + \frac{1}{n} - x^4 \\ &= \frac{2x^2}{n} + \frac{1}{n^2} + \frac{1}{n}\cdot\end{aligned}$$

La suite (h_n) converge simplement sur $\mathbb{R}$ vers la fonction nulle, mais la convergence ne peut pas être uniforme sur $\mathbb{R}$ puisque les h_n ne sont pas bornées.

Exercice 12.3

• Traitons d'abord le cas où $f = 0$.

Soit $\varepsilon > 0$. Choisissons $N \in \mathbb{N}$ tel que $k\dfrac{b-a}{N} \leqslant \varepsilon$, et considérons la subdivision de $[a\,;b]$ formée par les points $a_i = a + \dfrac{i(b-a)}{N}$ pour $0 \leqslant i \leqslant N$.

La convergence simple de la suite $(f_n)_{n\in\mathbb{N}}$ vers 0 en a_i implique que pour tout $i \in [\![0\,;N]\!]$ il existe $n_i \in \mathbb{N}$ tel que :

$$\forall\, n \geqslant n_i,\ \|f_n(a_i)\|_F \leqslant \varepsilon\,.$$

En posant $N' = \max\limits_{0\leqslant i\leqslant N} n_i$, on aura donc :

$$\forall\, i \in [\![0\,;N]\!],\ \forall\, n \geqslant N',\ \|f_n(a_i)\|_F \leqslant \varepsilon\,.$$

Tout $x \in [a\,;b]$ se trouve dans un intervalle de la forme $[a_i\,;a_{i+1}[$, donc pour $n \geqslant N'$ on aura grâce à l'inégalité triangulaire :

$$\begin{aligned}\|f_n(x)\|_F &\leqslant \|f_n(x)-f_n(a_i)\|_F + \|f_n(a_i)\|_F\\ &\leqslant k\,|x-a_i| + \varepsilon\\ &\leqslant k\,|a_{i+1}-a_i| + \varepsilon\\ &\leqslant k\frac{b-a}{N} + \varepsilon \leqslant 2\varepsilon\,,\end{aligned}$$

d'où $\|f_n\|_\infty \leqslant 2\varepsilon$.

Cela prouve la convergence uniforme de la suite $(f_n)_{n\in\mathbb{N}}$ vers la fonction nulle.

• Dans le cas général, on montre d'abord que la limite f est encore k-lipschitzienne : en effet, si x, $y \in [a\,;b]$, on a :

$$\forall\, n \in \mathbb{N},\ \|f_n(x) - f_n(y)\|_F \leqslant k\,|x-y|\ ,$$

donc par passage à la limite quand $n \to +\infty$, on a

$$\|f(x) - f(y)\|_F \leqslant k\,|x-y|\ .$$

Il en résulte que les fonctions $g_n = f_n - f$ sont $2k$-lipschitziennes. La suite $(g_n)_{n\in\mathbb{N}}$ converge simplement vers la fonction nulle donc, d'après le résultat précédent, la convergence est uniforme.

Exercice 12.4

• Puisque la fonction sin est bornée, on a pour tout $x > 0$: $\lim\limits_{n\to+\infty} f_n(x) = 0$. La suite $(f_n)_{n\in\mathbb{N}^*}$ converge donc simplement vers la fonction nulle sur $\mathbb{R}_+^*$.

• On connaît l'inégalité $|\sin x| \leqslant |x|$ valable pour tout x réel. On a donc

$$\forall\, x > 0,\ |f_n(x)| \leqslant \frac{nx}{n\sqrt{x}} = \sqrt{x}\,,$$

d'où :

$$\forall\, x \in \left]0\,;\frac{1}{n}\right],\ |f_n(x)| \leqslant \frac{1}{\sqrt{n}}\cdot$$

Puis, pour $x \in \left[\frac{1}{n}\,;+\infty\right[$, on a :

$$|f_n(x)| \leqslant \frac{|\sin nx|}{n\sqrt{\frac{1}{n}}} \leqslant \frac{1}{\sqrt{n}}\cdot$$

Finalement, $\|f_n\|_\infty^{\mathbb{R}_+^*} \leqslant \frac{1}{\sqrt{n}}$, c'est-à-dire que la suite $(f_n)_{n\in\mathbb{N}^*}$ converge uniformément vers la fonction nulle sur $\mathbb{R}_+^*$.

Exercice 12.5

• Étudions d'abord la convergence simple. La définition de f_n fait intervenir un intervalle qui dépend de n. Mais pour tout $x \in \mathbb{R}_+$ fixé, il existe un entier n_0 tel que $x \in [0\,;n_0[$ donc pour $n \geqslant n_0$ on aura :

$$f_n(x) = \left(1-\frac{x}{n}\right)^n = \mathrm{e}^{n\ln\left(1-\frac{x}{n}\right)}\,,$$

et puisque $\ln\left(1-\frac{x}{n}\right) \underset{n\to+\infty}{\sim} -\frac{x}{n}$, on obtient $\lim\limits_{n\to+\infty} f_n(x) = \mathrm{e}^{-x}$.

La suite $(f_n)_{n\in\mathbb{N}^*}$ converge donc simplement sur $\mathbb{R}_+$ vers $f : x \mapsto \mathrm{e}^{-x}$.

• Pour étudier la convergence uniforme, nous allons chercher $\|f_n - f\|_\infty$. Pour cela, posons $g_n = f_n - f$.

Pour $x \geqslant n$ on a $g_n(x) = -\mathrm{e}^{-x}$ et pour $x \in [0\,;n[$ on a

$$g_n(x) = \left(1-\frac{x}{n}\right)^n - \mathrm{e}^{-x}\,.$$

g_n est de classe $\mathscr{C}^1$ sur cet intervalle et

$$\forall\, x \in [0\,;n[,\ g_n'(x) = \mathrm{e}^{-x} - \left(1-\frac{x}{n}\right)^{n-1}\,.$$

Donc :

$$\begin{aligned}g_n'(x) \geqslant 0 &\iff \mathrm{e}^{-x} \geqslant \left(1-\frac{x}{n}\right)^{n-1}\\ &\iff -x \geqslant (n-1)\ln\left(1-\frac{x}{n}\right).\end{aligned}$$

On est donc amené à considérer la fonction φ définie pour $x \in [0\,;n[$ par :

$$\varphi(x) = (n-1)\ln\left(1-\frac{x}{n}\right) + x\,.$$

φ est de classe $\mathscr{C}^1$ et pour $x \in [0\,;n[$, $\varphi'(x) = \frac{x-1}{x-n}$ donc on a le tableau de variations suivant :

x	0		1		n
$\varphi'(x)$		$+$	0	$-$	
$\varphi(x)$	0	↗		↘	$-\infty$

Ce tableau montre qu'il existe un et un seul réel $\alpha_n > 1$ où φ s'annule et change de signe, et puisque le signe de g_n' sur $[0\,;n[$ est

l'opposé du signe de φ on aura le tableau de variations suivant (où l'on a complété g_n sur $[n\,;+\infty[$) :

x	0		α_n		n		$+\infty$
$g_n'(x)$		$-$	0	$+$		$+$	
$g_n(x)$	0	$\searrow$	$g_n(\alpha_n)$	$\nearrow$	$-\mathrm{e}^{-n}$	$\nearrow$	0

g_n étant à valeurs négatives, $\|g_n\|_\infty = |g_n(\alpha_n)|$. Or :

$$|g_n(\alpha_n)| = \mathrm{e}^{-\alpha_n} - \left(1-\frac{\alpha_n}{n}\right)^n ,$$

et puisque α_n est solution de l'équation $g_n'(x)=0$ on a :

$$\mathrm{e}^{-\alpha_n} = \left(1-\frac{\alpha_n}{n}\right)^{n-1}$$

donc :

$$\begin{aligned}|g_n(\alpha_n)| &= \mathrm{e}^{-\alpha_n} - \mathrm{e}^{-\alpha_n}\left(1-\frac{\alpha_n}{n}\right)\\ &= \frac{\alpha_n \mathrm{e}^{-\alpha_n}}{n}\,.\end{aligned}$$

Or une étude de fonction rapide permet de montrer que :

$$\forall x\in\mathbb{R}_+,\ x\mathrm{e}^{-x} \leqslant \frac{1}{\mathrm{e}}\,,$$

donc finalement :

$$\|f_n - f\|_\infty^{\mathbb{R}_+} \leqslant \frac{1}{n\mathrm{e}}\,.$$

Conclusion : la suite $(f_n)_{n\in\mathbb{N}^*}$ converge uniformément vers f sur $\mathbb{R}_+$.

Exercice 12.6

1) On distingue deux cas.

- Si $x=1$, on a $u_n(1)=0$ pour tout $n\in\mathbb{N}^*$, donc la série $\sum u_n(1)$ converge.
- Pour $x\in[0\,;1[$, la suite (a_n) étant positive et décroissante, on a :

  $$\forall x\in[0\,;1[,\ 0\leqslant u_n(x)\leqslant a_0 x^n(1-x)\,,$$

  et puisque la série géométrique de terme général $a_0x^n(1-x)$ et de raison x est convergente, on en déduit par comparaison de séries à termes positifs qu'il en est de même de la série $\sum u_n(x)$.

En conclusion : la série $\sum u_n$ converge simplement sur $[0\,;1]$.

2) Soit $f_n : x\mapsto x^n(1-x)$. Une étude rapide des variations de f_n montre qu'elle admet sur $[0\,;1]$ un maximum, atteint en $\frac{n}{n+1}$. On a donc :

$$\begin{aligned}\|f_n\|_\infty &= \left(\frac{n}{n+1}\right)^n\left(1-\frac{n}{n+1}\right)\\ &= \frac{1}{n+1}\left(1+\frac{1}{n}\right)^{-n}\\ &= \frac{1}{n+1}\mathrm{e}^{-n\ln\left(1+\frac{1}{n}\right)}\,,\end{aligned}$$

donc $\|f_n\|_\infty \underset{n\to+\infty}{\sim} \frac{1}{n\mathrm{e}}$ puis :

$$\|u_n\|_\infty \underset{n\to+\infty}{\sim} \frac{1}{\mathrm{e}}\,\frac{a_n}{n}\,.$$

Il en résulte, par comparaison de séries à termes positifs, que la série $\sum u_n$ converge normalement si et seulement si la série $\sum \frac{a_n}{n}$ converge.

3) Notons $R_n(x) = \sum_{k=n+1}^{+\infty} a_k x^k(1-x)$ le reste d'ordre n de la série.
On a $R_n(1)=0$ et, pour $x\in[0\,;1[$, la suite a étant décroissante positive :

$$\begin{aligned}0\leqslant R_n(x) &\leqslant a_{n+1}(1-x)\sum_{k=n+1}^{+\infty} x^k\\ &\leqslant a_{n+1}x^{n+1}\leqslant a_{n+1}\,.\end{aligned}$$

On en déduit $\|R_n\|_\infty^{[0;1]}\leqslant a_{n+1}$.

- Si la suite a converge vers 0, on a $\lim\limits_{n\to+\infty}\|R_n\|_\infty^{[0;1]} = 0$, donc la série converge uniformément sur $[0\,;1]$.
- Sinon, la suite a étant décroissante minorée converge vers un réel $\ell>0$. On a alors, pour tout $x\in[0\,;1[$:

  $$R_n(x)\geqslant \ell(1-x)\sum_{k=n+1}^{+\infty} x^k = \ell x^{n+1}\,,$$

  donc $\|R_n\|_\infty^{[0;1[}\geqslant\ell$ et puisque $R_n(1)=0$ on a aussi $\|R_n\|_\infty^{[0;1]}\geqslant\ell$.

  Dans ce cas, la convergence n'est pas uniforme sur $[0\,;1]$.

Cela établit l'équivalence demandée.

Exercice 12.7

1) Pour $x \in]0\,;1[$ on obtient directement par sommation d'une série géométrique :

$$\sum_{n=0}^{+\infty} u_n(x) = -\frac{x^2 \ln x}{1+x^2}\,,$$

cette relation restant valable pour $x = 0$ (en prolongeant) et $x = 1$.

2) Les hypothèses du critère spécial des séries alternées sont vérifiées (facile) et donc on a la majoration du reste :

$$\begin{aligned}|R_n(x)| &= \left|\sum_{k=n+1}^{+\infty} (-1)^{k+2} x^{2k+2} \ln x\right| \\ &\leqslant x^{2(n+2)} |\ln x|\,.\end{aligned}$$

L'étude de $\varphi\colon x \mapsto x^{2(n+2)}|\ln x|$ donne facilement :

$$\forall\, x \in [0\,;1],\ x^{2(n+2)}|\ln x| \leqslant \frac{1}{2(n+2)\mathrm{e}}\,,$$

donc $\|R_n\|_\infty \leqslant \dfrac{1}{2(n+2)\mathrm{e}} \underset{n\to+\infty}{\longrightarrow} 0\,.$

Cela démontre la convergence uniforme de la série sur $[0\,;1]$.

3) L'existence de l'intégrale est facile à vérifier. On a alors :

$$\begin{aligned}\int_0^1 \frac{\ln x}{1+x^2}\mathrm{d}x &= \int_0^1 \ln x \mathrm{d}x - \int_0^1 \frac{x^2 \ln x}{1+x^2}\mathrm{d}x \\ &= -1 + \int_0^1 \left(\sum_{n=0}^{+\infty} u_n(x)\right)\mathrm{d}x\,,\end{aligned}$$

et on peut calculer la dernière intégrale directement par intégration terme à terme puisqu'il y a convergence uniforme sur le segment $[0\,;1]$.

Une intégration par parties donne :

$$\begin{aligned}\int_0^1 x^{2n+2}\ln x &= \left[\frac{x^{2n+3}\ln x}{2n+3}\right]_0^1 \\ &\quad - \frac{1}{2n+3}\int_0^1 x^{2n+2}\,\mathrm{d}x \\ &= -\frac{1}{(2n+3)^2}\end{aligned}$$

d'où : $\displaystyle\int_0^1 \frac{\ln x}{1+x^2}\mathrm{d}x = -1 + \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+3)^2}\,,$

puis le résultat demandé.

Exercice 12.8

1) Pour tout entier $n \geqslant 1$ et x réel, notons $f_n(x) = \dfrac{1}{n^2x+n}$.

Pour que la série $\displaystyle\sum_{n\geqslant 1} f_n(x)$ existe, il faut déjà que $f_n(x)$ existe pour tout $n \geqslant 1$, ce qui implique $n^2x + n = n(nx+1) \neq 0$, soit $x \in \mathbb{R} \setminus \left\{-\dfrac{1}{n},\, n \in \mathbb{N}^*\right\}$.

Supposons donc cette condition vérifiée.

- Si $x = 0$, $f_n(0) = \dfrac{1}{n}$ donc la série $\displaystyle\sum_{n\geqslant 1} f_n(0)$ diverge (série harmonique).
- Si $x \neq 0$, alors $f_n(x) \underset{n\to+\infty}{\sim} \dfrac{1}{n^2 x}$. Puisque la série de terme général $\dfrac{1}{n^2}$ converge (série de Riemann), il en est de même de la série de terme général $f_n(x)$.

En conclusion : le domaine de définition de S est $\mathbb{R}^* \setminus \left\{-\dfrac{1}{n},\, n \in \mathbb{N}^*\right\}$.

2) On remarque déjà que, d'après ce qui précède, S est bien définie sur $\mathbb{R}_+^*$.

Puisque l'énoncé demande de montrer que S est de classe $\mathscr{C}^1$, on appliquera directement le théorème correspondant. On en vérifie soigneusement les hypothèses.

- Les fonctions $f_n : x \mapsto \dfrac{1}{n^2x+n}$ sont de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$.
- La série de fonctions $\displaystyle\sum_{n\geqslant 1} f_n$ converge simplement sur $\mathbb{R}_+^*$, c'est-à-dire que la série $\displaystyle\sum_{n\geqslant 1} f_n(x)$ converge pour tout $x > 0$: cela a déjà été fait.
- La dernière hypothèse à vérifier est la convergence uniforme, ou uniforme locale, de la série $\displaystyle\sum_{n\geqslant 1} f_n'$, c'est-à-dire de la série de fonctions $\displaystyle\sum_{n\geqslant 1} \frac{-n^2}{(n^2x+n)^2}$.

  Il est naturel de chercher si cette série converge normalement. Or, lorsque x tend vers 0, $f_n'(x)$ tend vers -1, donc $\|f_n'\|_\infty = \sup\limits_{x>0} |f_n'(x)| \geqslant 1$ et la série $\displaystyle\sum_{n\geqslant 1} \|f_n'\|_\infty$ diverge donc grossièrement.

Il n'y a donc pas convergence normale sur $\mathbb{R}_+^*$ tout entier. On peut donc essayer de considérer un ensemble plus petit.

Soit donc $a > 0$. Si $x \geqslant a$,

$$(n^2x+n)^2 \geqslant (n^2a+n)^2 \geqslant (n^2a)^2$$

donc $|f_n'(x)| \leqslant \dfrac{1}{n^2a^2}$.

Ainsi,

$$\left\|f_n'\right\|_\infty^{[a;+\infty[} = \sup_{x\geqslant a} \left|f_n'(x)\right| \leqslant \frac{1}{n^2a^2},$$

et, par comparaison de séries à termes positifs, la série $\displaystyle\sum_{n\geqslant 1} \left\|f_n'\right\|_\infty^{[a;+\infty[}$ converge.

La série de fonctions $\displaystyle\sum_{n\geqslant 1} f_n'$ converge donc normalement, donc uniformément, sur $[a\,;+\infty[$. Il résulte du théorème de dérivation d'une série de fonctions que S est de classe $\mathscr{C}^1$ sur tout intervalle de la forme $[a\,;+\infty[$ avec $a > 0$, et, les notions de continuité et de dérivabilité étant des notions locales, que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$.

3) Commençons par l'étude de S en $+\infty$.

On remarque d'abord que, pour tout $n \geqslant 1$, $\lim\limits_{x\to+\infty} f_n(x) = 0$. Puisque la série converge uniformément sur tout intervalle $[a\,;+\infty[$, le théorème de la double limite s'applique :

$$\lim_{x\to+\infty} S(x) = \sum_{n=1}^{+\infty} \lim_{x\to+\infty} f_n(x) = 0.$$

✓ Pour déterminer un équivalent, il faut un peu d'intuition.

On peut réfléchir ainsi (attention, ce n'est pas une démonstration rigoureuse !) : quand $x \to +\infty$, $f_n(x)$ est peu différent de $\dfrac{1}{n^2x}$ donc $S(x)$ devrait être peu différent de $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2x} = \frac{1}{x}\left(\sum_{n=1}^{+\infty} \frac{1}{n^2}\right)\ldots$

Ce type d'approche « à la physicienne » (et qui va se révéler juste !) fait penser à considérer $xS(x)$.

On écrit : $xS(x) = \displaystyle\sum_{n=1}^{+\infty} g_n(x)$, avec $g_n(x) = \dfrac{x}{n^2x+n}$. Pour tout $x > 0$,

$$0 \leqslant g_n(x) \leqslant \frac{x}{n^2x} = \frac{1}{n^2},$$

donc $\|g_n\|_\infty \leqslant \dfrac{1}{n^2}$.

Il en résulte que la série $\displaystyle\sum_{n\geqslant 1} g_n$ converge normalement, donc uniformément, sur $\mathbb{R}_+^*$. Puisque $\lim\limits_{x\to+\infty} g_n(x) = \dfrac{1}{n^2}$, on peut encore appliquer le théorème de la double limite :

$$\lim_{x\to+\infty} xS(x) = \sum_{n=1}^{+\infty} \lim_{x\to+\infty} g_n(x)$$
$$= \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

ce qui montre que $S(x) \underset{x\to+\infty}{\sim} \dfrac{\pi^2}{6x}$.

Autre méthode possible

Pour obtenir un équivalent de $S(x)$ on peut aussi essayer d'encadrer $S(x)$ par deux expressions équivalentes entre elles (ce sera également la méthode utilisée dans la question suivante).

On remarque ici que l'on a :

$\forall n \in \mathbb{N}^*, \forall x > 0,$

$$\frac{1}{n^2x+n^2} \leqslant \frac{1}{n^2x+n} \leqslant \frac{1}{n^2x},$$

et en additionnant ces inégalités pour n variant de 1 à $+\infty$, ce qui est possible car toutes les séries écrites convergent, on obtient, pour tout $x > 0$:

$$\frac{1}{x+1}\sum_{n=1}^{+\infty}\frac{1}{n^2} \leqslant S(x) \leqslant \frac{1}{x}\sum_{n=1}^{+\infty}\frac{1}{n^2}$$

soit

$$\forall x > 0,\ \frac{\pi^2}{6(x+1)} \leqslant S(x) \leqslant \frac{\pi^2}{6x},$$

ce qui permet de retrouver le résultat précédent.

4) Pour l'étude de S en 0^+, on va utiliser la méthode de comparaison série-intégrale, puisqu'on a remarqué que, à $x > 0$ fixé, la suite $n \mapsto \dfrac{1}{n^2x+n}$ est décroissante.

✓ Il faut apprendre à reconnaître des situations où la méthode de comparaison série-intégrale s'applique.

Soit donc $x > 0$ fixé. On considère la fonction $\varphi\colon t \mapsto \dfrac{1}{t^2x+t}$; φ est continue et décroissante sur $\mathbb{R}_+^*$. On aura donc :

$$\forall n \geqslant 1,\ \forall t \in [n\,;n+1],\ \varphi(t) \leqslant \varphi(n)$$

et

$$\forall n \geqslant 2,\ \forall t \in [n-1\,;n],\ \varphi(n) \leqslant \varphi(t).$$

ce qui donne l'encadrement, pour $n \geqslant 2$:

$$\int_n^{n+1} \varphi(t)\,\mathrm{d}t \leqslant \varphi(n) = f_n(x) \leqslant \int_{n-1}^{n} \varphi(t)\,\mathrm{d}t$$

(faire particulièrement attention aux indices : φ n'est pas définie en 0, et l'intégrale $\int_{n-1}^{n} \varphi(t)\,\mathrm{d}t$ n'aurait pas de sens pour $n = 1$!).

✓ Il est *important* de savoir retrouver rapidement l'encadrement ci-dessus ; un petit graphique est bien utile pour cela, voir **[S11.7]** page 122.

Soit alors N un entier $\geqslant 2$. En additionnant les inégalités $\int_n^{n+1} \varphi(t)\,\mathrm{d}t \leqslant f_n(x)$ pour n variant de 1 à N et les inégalités $f_n(x) \leqslant \int_{n-1}^{n} \varphi(t)\,\mathrm{d}t$ pour n variant de 2 à N, on obtient, sachant que $f_1(x) = \dfrac{1}{x+1}$:

$$\int_1^{N+1} \varphi(t)\,\mathrm{d}t \leqslant \sum_{n=1}^{N} f_n(x) \leqslant \int_1^{N} \varphi(t)\,\mathrm{d}t + \frac{1}{x+1}\cdot$$

Or l'intégrale $\int_1^{+\infty} \varphi(t)\,\mathrm{d}t$ converge : cela est une conséquence du théorème de comparaison série-intégrale, ou cela peut être montré directement en remarquant que $\varphi(t) \underset{t\to+\infty}{\sim} \dfrac{1}{xt^2}$. On peut donc faire tendre N vers $+\infty$ dans l'encadrement précédent, ce qui conduit à :

$$\int_1^{+\infty} \varphi(t)\,\mathrm{d}t \leqslant S(x) \leqslant \int_1^{+\infty} \varphi(t)\,\mathrm{d}t + \frac{1}{x+1}$$

pour tout $x > 0$.

Pour calculer l'intégrale :

$$\int_1^{+\infty} \varphi(t)\,\mathrm{d}t = \int_1^{+\infty} \frac{\mathrm{d}t}{t(1+tx)}\,,$$

on fait une décomposition en éléments simples :

$$\frac{1}{t(1+tx)} = \frac{1}{t} - \frac{x}{1+tx}\cdot$$

Donc, pour tout réel $A \geqslant 1$:

$$\begin{aligned}\int_1^{A} \varphi(t)\,\mathrm{d}t &= \int_1^{A} \left(\frac{1}{t} - \frac{x}{1+tx}\right)\mathrm{d}t \\ &= \Big[\ln t - \ln(1+tx)\Big]_{t=1}^{t=A} \\ &= \ln\left(\frac{A(1+x)}{1+Ax}\right),\end{aligned}$$

puis, quand A tend vers $+\infty$:

$$\int_1^{+\infty} \varphi(t)\,\mathrm{d}t = \ln\left(\frac{1+x}{x}\right)\cdot$$

On a donc finalement l'encadrement :

$$\ln\left(\frac{1+x}{x}\right) \leqslant S(x) \leqslant \ln\left(\frac{1+x}{x}\right) + \frac{1}{x+1}$$

pour tout $x > 0$.

On constate que les termes figurant dans les membres gauche et droit de cette inégalité sont tous deux équivalents à $-\ln x$ quand x tend vers 0^+. Pour obtenir un équivalent de $S(x)$, on divise donc cette inégalité par $-\ln x$, qui est strictement positif pour $x \in]0\,;1[$:

$$\begin{aligned}&\forall\, x \in]0\,;1[,\\ &1 - \frac{\ln(x+1)}{\ln x} \leqslant \frac{S(x)}{-\ln x} \\ &\qquad\leqslant 1 - \frac{\ln(x+1)}{\ln x} - \frac{1}{(x+1)\ln x}\cdot\end{aligned}$$

Puisque les membres gauche et droit de cette inégalité tendent vers 1 quand x tend vers 0^+, on a donc $\lim\limits_{x\to0^+} \dfrac{S(x)}{-\ln x} = 1$ soit :

$$S(x) \underset{x\to0^+}{\sim} -\ln x\,.$$

Exercice 12.9

• Démontrons déjà que f est bien définie sur $\mathbb{R}_+^*$:

– les fonctions $x \mapsto \dfrac{1}{\operatorname{sh} nx}$ pour $n \geqslant 1$ sont bien définies sur $\mathbb{R}_+^*$;

– puisque $\operatorname{sh} X \underset{X\to+\infty}{\sim} \dfrac{\mathrm{e}^X}{2}$, on a

$$\forall\, x > 0,\ \frac{1}{\operatorname{sh} nx} \underset{n\to+\infty}{\sim} 2\mathrm{e}^{-nx}$$

et puisque e^{-nx} est le terme général d'une série géométrique de raison $\mathrm{e}^{-x} \in]0\,;1[$, la série proposée converge d'après les critères de comparaison.

• *Étude en* $+\infty$

Pour $X \geqslant 1$ on a facilement $\frac{1}{2}\mathrm{e}^X \geqslant \mathrm{e}^{-X}$, donc pour $n \geqslant 1$ et $x \geqslant 1$:

$$\mathrm{e}^{nx} - \mathrm{e}^{-nx} \geqslant \frac{1}{2}\mathrm{e}^{nx}.$$

On en déduit :

$$\frac{1}{\operatorname{sh} nx} = \frac{2}{\mathrm{e}^{nx} - \mathrm{e}^{-nx}} \leqslant 4\mathrm{e}^{-nx},$$

puis, pour $x \geqslant 1$:

$$\begin{aligned} 0 \leqslant f(x) - \frac{1}{\operatorname{sh} x} &= \sum_{n=2}^{+\infty} \frac{1}{\operatorname{sh} nx} \\ &\leqslant 4\sum_{n=2}^{+\infty} \mathrm{e}^{-nx} = \frac{4\mathrm{e}^{-2x}}{1-\mathrm{e}^{-x}}. \end{aligned}$$

Toujours à l'aide de l'équivalent $\operatorname{sh} x \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^x}{2}$, on déduit de cette inégalité que :

$$\lim_{x\to+\infty} \operatorname{sh} x \left(f(x) - \frac{1}{\operatorname{sh} x}\right) = 0,$$

c'est-à-dire que

$$f(x) - \frac{1}{\operatorname{sh} x} \underset{x\to+\infty}{=} o\left(\frac{1}{\operatorname{sh} x}\right)$$

ou encore :

$$f(x) \underset{x\to+\infty}{\sim} \frac{1}{\operatorname{sh} x} \underset{x\to+\infty}{\sim} 2\mathrm{e}^{-x}.$$

• *Étude en* 0^+

Lorsque x tend vers 0 on a :

$$\frac{1}{\operatorname{sh} nx} = \frac{2\mathrm{e}^{-nx}}{1-\mathrm{e}^{-2nx}} \underset{x\to0}{\sim} \frac{\mathrm{e}^{-nx}}{nx}$$

et on voudrait pouvoir additionner ces équivalents. Pour cela, il faut essayer de majorer la différence.

On a donc, plus précisément :

$$\begin{aligned} \frac{1}{\operatorname{sh} nx} - \frac{\mathrm{e}^{-nx}}{nx} &= \frac{2\mathrm{e}^{-nx}}{1-\mathrm{e}^{-2nx}} - \frac{2\mathrm{e}^{-nx}}{2nx} \\ &= 2\mathrm{e}^{-nx}\left(\frac{1}{1-\mathrm{e}^{-2nx}} - \frac{1}{2nx}\right) \\ &= 2\mathrm{e}^{-nx}\varphi(2nx) \end{aligned}$$

où l'on a posé $\varphi(t) = \frac{1}{1-\mathrm{e}^{-t}} - \frac{1}{t}$.

Cette fonction φ est continue sur $\mathbb{R}_+^*$; de plus,

$$\begin{aligned} \lim_{t\to0} \varphi(t) &= \lim_{t\to0} \frac{\mathrm{e}^{-t} - (1-t)}{t(1-\mathrm{e}^{-t})} \\ &= \lim_{t\to0} \frac{t^2/2}{t^2} = \frac{1}{2} \end{aligned}$$

(à l'aide d'un développement limité) et

$$\lim_{t\to+\infty} \varphi(t) = 1.$$

Il en résulte que φ est bornée sur $\mathbb{R}_+$, d'où, pour tout $x > 0$ et tout $n \in \mathbb{N}^*$:

$$\left|\frac{1}{\operatorname{sh} nx} - \frac{\mathrm{e}^{-nx}}{nx}\right| \leqslant 2\|\varphi\|_\infty \mathrm{e}^{-nx}.$$

En sommant et par l'inégalité triangulaire on en déduit, pour tout $x > 0$:

$$\left|\sum_{n=1}^{+\infty} \frac{1}{\operatorname{sh} nx} - \frac{\mathrm{e}^{-nx}}{nx}\right| \leqslant 2\|\varphi\|_\infty \sum_{n=1}^{+\infty} \mathrm{e}^{-nx}$$

soit

$$\left|f(x) - \sum_{n=1}^{+\infty} \frac{\mathrm{e}^{-nx}}{nx}\right| \leqslant 2\|\varphi\|_\infty \frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}.$$

Or on connaît le développement en série entière :

$$\forall\, x \in]-1\,;1[,\ \ln(1-x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n},$$

donc pour tout $x > 0$:

$$\left|f(x) + \frac{\ln(1-\mathrm{e}^{-x})}{x}\right| \leqslant 2\|\varphi\|_\infty \frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}}$$

et puisque $\frac{\mathrm{e}^{-x}}{1-\mathrm{e}^{-x}} \underset{x\to0}{\sim} \frac{1}{x}$, on en déduit :

$$f(x) = -\frac{\ln(1-\mathrm{e}^{-x})}{x} + O\left(\frac{1}{x}\right).$$

Enfin, le développement limité :

$$\ln(1-\mathrm{e}^{-x}) = \ln\left(x + O\left(x^2\right)\right) = \ln x + o(1)$$

permet de conclure :

$$f(x) \underset{x\to0^+}{=} -\frac{\ln x}{x} + O\left(\frac{1}{x}\right).$$

Exercice 12.10

1) • Pour $x = 0$ on a $u_n(0) = (-1)^n$ donc la série $\sum u_n(0)$ diverge grossièrement.

Pour $x > 0$, il est facile de vérifier que la série $\sum u_n(x)$ vérifie le critère spécial des séries alternées, donc converge.

En conclusion, la série $\sum u_n$ converge simplement sur $\mathbb{R}_+^*$.

• En notant R_n le reste d'ordre n de cette série, on sait que :

$$\forall\, x > 0,\ |R_n(x)| \leqslant \frac{1}{\sqrt{(n+1)x+1}},$$

donc pour tout réel $a > 0$ on a :

$$\|R_n\|_\infty^{[a;+\infty[} \leqslant \frac{1}{\sqrt{(n+1)a+1}} \underset{n\to+\infty}{\longrightarrow} 0\,.$$

Il y a donc convergence uniforme sur tout intervalle $[a\,;+\infty[$ (ce qui permettrait de prouver la continuité de S sur $\mathbb{R}_+^*$, mais ce n'était pas demandé).

2) • Puisqu'il y a convergence uniforme au voisinage de $+\infty$, le théorème de la double limite s'applique et donne directement

$$\lim_{x\to+\infty} S(x) = 0\,.$$

• Puisque $u_n(x) \underset{x\to+\infty}{\sim} \dfrac{(-1)^n}{\sqrt{nx}}$, on voudrait pouvoir additionner ces équivalents.
Pour cela on calcule, pour tout $n \in \mathbb{N}^*$ et tout $x > 0$:

$$\begin{aligned}\frac{1}{\sqrt{nx}} - \frac{1}{\sqrt{nx+1}} &= \frac{\sqrt{nx+1}-\sqrt{nx}}{\sqrt{nx}\sqrt{nx+1}} \\ &= \frac{1}{\sqrt{nx}\sqrt{nx+1}\left(\sqrt{nx}+\sqrt{nx+1}\right)} \\ &\leqslant \frac{1}{\sqrt{nx}\sqrt{nx}\left(\sqrt{nx}+\sqrt{nx}\right)} = \frac{1}{2(nx)^{3/2}}\,.\end{aligned}$$

Ainsi, $\left|u_n(x) - \dfrac{(-1)^n}{\sqrt{nx}}\right| \leqslant \dfrac{1}{2(nx)^{3/2}}$.

Puisque la série de terme général $\dfrac{(-1)^n}{\sqrt{nx}}$ et celle de terme général $\dfrac{1}{2(nx)^{3/2}}$ convergent, on obtient, en sommant et grâce à l'inégalité triangulaire :

$$\begin{aligned}&\left|S(x) - \sum_{n=1}^{+\infty} \frac{(-1)^n}{\sqrt{nx}}\right| \\ &= \left|\sum_{n=1}^{+\infty}\left(u_n(x) - \frac{(-1)^n}{\sqrt{nx}}\right)\right| \\ &\leqslant \sum_{n=1}^{+\infty}\left|u_n(x) - \frac{(-1)^n}{\sqrt{nx}}\right| \leqslant \sum_{n=1}^{+\infty}\frac{1}{2(nx)^{3/2}} \\ &\leqslant \frac{1}{2x^{3/2}}\left(\sum_{n=1}^{+\infty}\frac{1}{n^{3/2}}\right).\end{aligned}$$

En notant $\ell = \displaystyle\sum_{n=1}^{+\infty} \frac{(-1)^n}{\sqrt{n}}$ on a finalement :

$$S(x) - \frac{\ell}{\sqrt{x}} = O\left(\frac{1}{x^{3/2}}\right),$$

et le développement asymptotique demandé est :

$$S(x) \underset{x\to0^+}{=} \frac{\ell}{\sqrt{x}} + O\left(\frac{1}{x^{3/2}}\right).$$

Exercice 13.1

• Puisque $\dfrac{1}{(-1)^n+\sqrt{n}} \underset{n\to+\infty}{\sim} \dfrac{1}{\sqrt{n}}$, que $\lim\limits_{n\to+\infty} \dfrac{1}{\sqrt{n}} = 0$ et que $\sin x \underset{x\to0}{\sim} x$, on a $a_n \underset{n\to+\infty}{\sim} \dfrac{1}{\sqrt{n}}$.

La série entière $\sum a_n x^n$ a donc même rayon de convergence que la série entière $\sum \dfrac{x^n}{\sqrt{n}}$, c'est-à-dire 1 (par exemple par le critère de d'Alembert, puisque $\lim\limits_{n\to+\infty} \dfrac{\sqrt{n}}{\sqrt{n+1}} = 1$).

• Cela ne répond pas tout à fait à la question : il reste à examiner ce qui se passe pour $x = \pm1$.

– Pour $x = 1$: la série $\sum a_n$ diverge puisque $a_n \sim \dfrac{1}{\sqrt{n}}$ et que la série à termes positifs $\sum \dfrac{1}{\sqrt{n}}$ diverge (série de Riemann).

– Pour $x = -1$, il s'agit d'étudier la série de terme général $(-1)^n a_n$. Celui-ci n'étant pas de signe constant, un simple équivalent ne suffit pas ; nous allons donc utiliser un développement limité de a_n. On a d'abord :

$$\begin{aligned}\frac{1}{(-1)^n+\sqrt{n}} &= \frac{1}{\sqrt{n}\left(1+\dfrac{(-1)^n}{\sqrt{n}}\right)} \\ &= \frac{1}{\sqrt{n}}\left(1 - \frac{(-1)^n}{\sqrt{n}} + O\left(\frac{1}{n}\right)\right),\end{aligned}$$

puis :

$$\begin{aligned}a_n &= \sin\left(\frac{1}{\sqrt{n}} - \frac{(-1)^n}{n} + O\left(\frac{1}{n^{3/2}}\right)\right) \\ &= \frac{1}{\sqrt{n}} - \frac{(-1)^n}{n} + O\left(\frac{1}{n^{3/2}}\right).\end{aligned}$$

On en déduit que :

$$(-1)^n a_n = \frac{(-1)^n}{\sqrt{n}} - \frac{1}{n} + O\left(\frac{1}{n^{3/2}}\right).$$

Or la série de terme général $\dfrac{(-1)^n}{\sqrt{n}}$ converge (d'après le critère spécial sur les séries alternées), la série de terme général $\dfrac{1}{n}$ diverge (série harmonique),

et une série dont le terme général est un $O\left(\frac{1}{n^{3/2}}\right)$ est absolument convergente (d'après les règles de comparaison sur les séries à termes positifs).

La série de terme général $(-1)^n a_n$ est donc la somme d'une série divergente et de deux séries convergentes : elle diverge.

En conclusion, le domaine de définition de $\sum a_n x^n$ est $]-1\,;1[$.

Exercice 13.2

Puisque pour $n \in \mathbb{N}^*$ on a $\frac{1}{n} \leqslant n^{(-1)^n} \leqslant n$, le rayon de convergence de la série proposée est encadré par ceux des séries entières (connues) $\sum \frac{x^n}{n}$ et $\sum x^n$, donc vaut 1.

Notons $f(x)$ sa somme.

En séparant les termes d'indices pairs et impairs, puisque les deux séries ont aussi un rayon de convergence égal à 1 on aura, pour $x \in]-1\,;1[$:

$$f(x) = \sum_{n=0}^{+\infty} 2nx^{2n} + \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1}\,.$$

Il reste à exprimer ces deux séries entières à l'aide de fonctions connues.

• Puisque l'on peut dériver une série entière terme à terme sur son intervalle de convergence, on a, pour tout $x \in]-1\,;1[$:

$$\sum_{\cancel{n=0}\ n=1}^{+\infty} 2nx^{2n} = x\sum_{n=1}^{+\infty} 2nx^{2n-1}$$

$$= x\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=1}^{+\infty} x^{2n}\right)$$

$$= x\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^2}{1-x^2}\right) = \frac{2x^2}{(1-x^2)^2}\,.$$

• D'autre part, en utilisant ici l'intégration terme à terme d'une série entière, on a, pour tout $x \in]-1\,;1[$:

$$\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1} = \int_0^x \left(\sum_{n=0}^{+\infty} t^{2n}\right)\mathrm{d}t$$

$$= \int_0^x \frac{\mathrm{d}t}{1-t^2}$$

$$= \frac{1}{2}\left(\int_0^x \frac{\mathrm{d}t}{1-t} + \int_0^x \frac{\mathrm{d}t}{1+t}\right)$$

$$= \frac{1}{2}\ln\frac{1+x}{1-x}$$

(on aurait pu d'ailleurs obtenir ce résultat à l'aide du développement en série entière de la fonction ln).

En conclusion, pour tout $x \in]-1\,;1[$,

$$f(x) = \frac{2x^2}{(1-x^2)^2} + \frac{1}{2}\ln\frac{1+x}{1-x}\,.$$

Exercice 13.3

1) On utilisera ici la caractérisation suivante du rayon de convergence :

$$R = \sup\left\{|z|\,,\ z \in \mathbb{C}\ \middle|\ \lim_{n\to+\infty} a_n z^n = 0\right\}\,.$$

Notons R le rayon de convergence de la série $\sum a_n z^n$ et R' celui de la série entière $\sum a_n^2 z^n$.

• Soit $|z| < R$.

Puisque $\lim\limits_{n\to+\infty} a_n z^n = 0$, on a aussi $\lim\limits_{n\to+\infty} a_n^2 z^{2n} = 0$ donc $|z|^2 \leqslant R'$.

Cela étant vrai pour tout z tel que $|z| < R$, on en déduit $R^2 \leqslant R'$ (en faisant tendre $|z|$ vers R par exemple).

• Soit $|z| < \sqrt{R'}$.

Alors $|z|^2 < R'$ donc $\lim\limits_{n\to+\infty} \left|a_n^2 z^{2n}\right| = 0$ d'où $\lim\limits_{n\to+\infty} |a_n z^n| = 0$ (en prenant la racine du module dans $\mathbb{R}$). On en déduit $|z| \leqslant R$.

Cela étant valable pour tout z tel que $|z| < \sqrt{R'}$, on en déduit $\sqrt{R'} \leqslant R$

En conclusion :

$$R' = R^2\,.$$

2) Soit $r < R$, $r \neq 0$. On sait que la suite $(a_n r^n)$ est bornée donc il existe un réel M positif tel que :

$$\forall\, n \in \mathbb{N}, |a_n r^n| \leqslant M \text{ soit } |a_n| \leqslant \frac{M}{r^n}\,.$$

Pour tout $z \in \mathbb{C}$ on aura alors

$$\left|\frac{a_n}{n!} z^n\right| \leqslant \frac{M}{n!}\left|\frac{z}{r}\right|^n \quad (*)$$

Or on sait que pour tout x réel la série $\sum \frac{x^n}{n!}$ est convergente (de somme e^x), donc par comparaison de séries à termes réels positifs, on déduit de l'inégalité $(*)$ que la série $\sum \frac{a_n z^n}{n!}$ est absolument convergente pour tout $z \in \mathbb{C}$.

Le rayon de convergence de la série entière considérée est donc $+\infty$.

Exercice 13.4

• Puisque la série $\sum a_n$ diverge, le rayon de convergence de la série entière $\sum a_n z^n$ est inférieur ou égal à 1.

• On a, pour $n \geqslant 1$:

$$\frac{a_n}{s_n} = \frac{s_n - s_{n-1}}{s_n} = 1 - \frac{s_{n-1}}{s_n}\,,$$

donc $\lim\limits_{n\to+\infty} \dfrac{s_{n-1}}{s_n} = 1$. D'après le critère de d'Alembert, le rayon de convergence de la série entière $\sum s_n z^n$ est égal à 1.

• Il en découle que, puisque $a_n \leqslant s_n$, le rayon de convergence de la série entière $\sum a_n z^n$ est supérieur ou égal à 1.

• Finalement, le rayon de convergence des deux séries entières est égal à 1, et puisque $a_n = s_n - s_{n-1}$ pour $n \geqslant 1$ et $a_0 = s_0$, on a, pour tout z tel que $|z| < 1$:

$$\sum_{n=0}^{+\infty} a_n z^n = (1-z)\sum_{n=0}^{+\infty} s_n z^n\,.$$

Exercice 13.5

• D'après le cours, les fonctions $x \mapsto \operatorname{Arcsin} x$ et $x \mapsto (1-x^2)^{-1/2}$ sont développables en série entière sur $]-1\,;1[$. Leur produit est donc aussi développable en série entière, avec un rayon de convergence a priori $\geqslant 1$. f n'étant pas définie en 1, ce rayon de convergence ne peut être qu'égal à 1.

• Un calcul rapide montre que f vérifie l'équation différentielle :

$$\forall\, x \in\,]-1\,;1[,\ (1-x^2)f'(x) - xf(x) = 1\,.$$

f étant développable en série entière et impaire, on sait qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ de réels tels que :

$$\forall\, x \in\,]-1\,;1[,\ f(x) = \sum_{n=0}^{+\infty} a_n x^{2n+1}$$

d'où aussi :

$$\forall\, x \in\,]-1\,;1[,\ f'(x) = \sum_{n=0}^{+\infty} (2n+1)a_n x^{2n}$$

et en remplaçant dans l'équation différentielle on obtient :

$$\sum_{n=0}^{+\infty}(2n+1)a_n x^{2n} - \sum_{n=0}^{+\infty}(2n+1)a_n x^{2n+2} - \sum_{n=0}^{+\infty} a_n x^{2n+2} = 1\,,$$

puis, après un décalage d'indices dans la première série :

$$a_0 + \sum_{n=0}^{+\infty}\Big[(2n+3)a_{n+1} - (2n+1)a_n - a_n\Big]x^{2n+2} = 1\,.$$

Par unicité du développement en série entière on obtient :

$$a_0 = 1 \ \text{ et } \ \forall\, n \in \mathbb{N},\ a_{n+1} = \frac{2n+2}{2n+3}a_n\,.$$

Donc :

$$a_n = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdots\frac{2}{3} = \frac{[2n(2n-2)\cdots 2]^2}{(2n+1)!} = \frac{(2^n n!)^2}{(2n+1)!}\,.$$

On en conclut : $\forall\, x \in\,]-1\,;1[\,,$

$$f(x) = \sum_{n=0}^{+\infty} \frac{(2^n n!)^2}{(2n+1)!}\,x^{2n+1}\,.$$

✓ Notez comment a été utilisé le théorème sur le produit de Cauchy de deux séries entières dans cet exemple : non pour calculer les coefficients (cela aurait donné des formules horribles), mais seulement pour démontrer l'existence du développement en série entière et connaître son rayon de convergence.

Exercice 13.6

1) Nous allons calculer les intégrales :

$$I_{n,m} = \int_0^1 t^n(1-t)^m\,\mathrm{d}t\,,\ (n,m) \in \mathbb{N}^2\,,$$

par récurrence sur m.
Une intégration par parties simple donne, pour $m \geqslant 1$:

$$I_{n,m} = \left[\frac{t^{n+1}}{n+1}\cdot(1-t)^m\right]_0^1 - \int_0^1 \frac{t^{n+1}}{n+1}(-m)(1-t)^{m-1}\,\mathrm{d}t = \frac{m}{n+1}\,I_{n+1,m-1}\,.$$

On aura donc :

$$\begin{aligned} I_{n,m} &= \frac{m}{n+1}\, I_{n+1,m-1} \\ &= \frac{m}{n+1}\cdot\frac{m-1}{n+2}\, I_{n+2,m-2} = \cdots \\ &= \frac{m}{n+1}\cdot\frac{m-1}{n+2}\cdots\frac{1}{n+m}\, I_{n+m,0} \\ &= \frac{m!\,n!}{(n+m)!}\, I_{n+m,0}\,, \end{aligned}$$

et puisque

$$I_{n+m,0} = \int_0^1 t^{n+m}\,\mathrm{d}t = \frac{1}{n+m+1}\,,$$

on trouve :

$$I_{n,m} = \frac{m!\,n!}{(n+m+1)!}$$

et en particulier :

$$I_{n,n} = a_n = \frac{(n!)^2}{(2n+1)!}\cdot$$

2) Puisque $a_n \neq 0$, on peut utiliser le critère de d'Alembert :

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{(n+1)^2}{(2n+2)(2n+3)} \\ &= \frac{n+1}{2(2n+3)} \underset{n\to+\infty}{\longrightarrow} \frac{1}{4}\,. \end{aligned}$$

On en déduit le rayon de convergence de la série entière $\sum\limits_{n\geqslant 0} a_n x^n$ est $R = 4$.

3) Il y a ici principalement deux méthodes pour calculer la somme de la série entière. La seconde méthode semble moins directe compte tenu de l'énoncé, mais nous semble plus élégante, et généralisable à d'autres cas (voir aussi l'exercice **13.8**).

• *Première méthode : interversion série-intégrale*

Soit $x \in]-4\,;4[$.
Posons alors, pour tout $t \in [0\,;1]$, $f_n(t) = t^n(1-t)^n x^n$.
Puisque $0 \leqslant t(1-t) \leqslant \frac{1}{4}$ pour tout $t \in [0\,;1]$ (sommet d'une parabole...), on a $|f_n(t)| \leqslant \left(\frac{|x|}{4}\right)^n$ pour tout $t \in [0\,;1]$.
Cette majoration étant indépendante de t, et puisque la série géométrique $\sum\limits_{n\in\mathbb{N}} \left(\frac{|x|}{4}\right)^n$ converge, cela implique que la série de fonctions $\sum\limits_{n\in\mathbb{N}} f_n(t)$ converge normalement, donc uniformément, sur $[0\,;1]$.

✓ Attention dans ce genre d'exercices : ne confondez pas x et t ! L'exercice concerne la série entière fonction de x, mais pour intervertir série et intégrale par rapport à t, on fixe x et on considère une série de fonctions de t.

La convergence de la série de fonctions (continues) $\sum\limits_{n\in\mathbb{N}} f_n(t)$ étant uniforme sur le *segment* $[0\,;1]$, on peut écrire :

$$\int_0^1 \left(\sum_{n=0}^{+\infty} f_n(t)\right)\mathrm{d}t = \sum_{n=0}^{+\infty}\left(\int_0^1 f_n(t)\,\mathrm{d}t\right),$$

soit, en notant f la somme de la série entière :

$$\begin{aligned} f(x) = \sum_{n=0}^{+\infty} a_n x^n &= \int_0^1 \sum_{n=0}^{+\infty} \big[t(t-1)x\big]^n\,\mathrm{d}t \\ &= \int_0^1 \frac{\mathrm{d}t}{1-t(1-t)x} = \int_0^1 \frac{\mathrm{d}t}{xt^2-xt+1}\cdot \end{aligned}$$

Il y a donc trois cas à distinguer (selon le signe du discriminant du trinôme en t figurant au dénominateur).

– Si $x = 0$, on a directement $f(0) = 1$.

– Si $x \in]0\,;4[$:

$$\begin{aligned} f(x) &= \int_0^1 \frac{\mathrm{d}t}{x\left(t-\frac{1}{2}\right)^2 + 1 - \frac{x}{4}} \\ &= \frac{1}{x}\int_0^1 \frac{\mathrm{d}t}{\left(t-\frac{1}{2}\right)^2 + \frac{4-x}{4x}} \\ &= \frac{1}{x}\sqrt{\frac{4x}{4-x}} \\ &\quad\times\left[\operatorname{Arctan}\left(\sqrt{\frac{4x}{4-x}}\left(t-\frac{1}{2}\right)\right)\right]_0^1 \\ &= \frac{4}{\sqrt{x(4-x)}}\operatorname{Arctan}\sqrt{\frac{x}{4-x}}\cdot \end{aligned}$$

– Si $x \in]-4\,;0[$:

$$\begin{aligned} f(x) &= \int_0^1 \frac{\mathrm{d}t}{x\left(t-\frac{1}{2}\right)^2 + 1 - \frac{x}{4}} \\ &= \frac{1}{-x}\int_0^1 \frac{\mathrm{d}t}{\frac{x-4}{4x} - \left(t-\frac{1}{2}\right)^2} \\ &= \frac{1}{-x}\sqrt{\frac{4x}{x-4}} \\ &\quad\times\left[\operatorname{Argth}\left(\sqrt{\frac{4x}{x-4}}\left(t-\frac{1}{2}\right)\right)\right]_0^1 \\ &= \frac{4}{\sqrt{x(x-4)}}\operatorname{Argth}\sqrt{\frac{x}{x-4}}, \end{aligned}$$

avec, pour $|y|<1$,

$$\operatorname{Argth}(y)=\frac{1}{2}\ln\left(\frac{1+y}{1-y}\right).$$

– Mais ce n'est pas fini ! il reste à examiner ce qui se passe pour $x=\pm 4$.

– Pour $x=4$: on a $a_n x^n=\dfrac{(n!)^2\,4^n}{(2n+1)!}$ et la formule de Stirling donne, lorsque $n\to+\infty$:

$$a_n\,4^n\sim\frac{\left(\frac{n}{\mathrm{e}}\right)^{2n}\cdot 2\pi n\cdot 4^n}{(2n+1)\left(\frac{2n}{\mathrm{e}}\right)^{2n}\sqrt{4\pi n}}\sim\frac{\sqrt{\pi}}{2\sqrt{n}}$$

et puisque la série de terme général $\dfrac{1}{\sqrt{n}}$, à termes positifs, diverge, il en est de même de la série de terme général $a_n\,4^n$.

La fonction f n'est donc pas définie pour $x=4$.

– Pour $x=-4$:

Lorsque $x\in[-4\,;0]$, on a vu que

$$\frac{\left|a_{n+1}x^{n+1}\right|}{|a_n x^n|}=\frac{n+1}{2(2n+3)}\,|x|<1\,.$$

La suite $\left(|a_n x^n|\right)$ est donc décroissante. Puisque $|a_n x^n|\leqslant|a_n\,4^n|$, le calcul d'équivalent ci-dessus montre que cette suite tend vers 0, donc la série de terme général $a_n x^n$ vérifie le critère spécial sur les séries alternées.

Par conséquent, son reste de rang n sera majoré en valeur absolue par $\left|a_{n+1}x^{n+1}\right|$, donc par $\left|a_{n+1}\,4^{n+1}\right|$, qui est indépendant de x et qui tend vers 0 quand $n\to+\infty$.

Ainsi, la série de terme général $a_n x^n$ converge en fait uniformément sur $[-4\,;0]$, ce qui implique que sa fonction somme f est continue sur $[-4\,;0]$. On aura donc

$$\begin{aligned}f(-4)&=\lim_{x\to-4^+}f(x)\\&=\lim_{x\to-4^+}\frac{4}{\sqrt{x^2-4x}}\operatorname{Argth}\sqrt{\frac{x}{x-4}}\\&=\frac{1}{\sqrt{2}}\operatorname{Argth}\frac{1}{\sqrt{2}}\cdot\end{aligned}$$

• *Seconde méthode : utilisation d'une équation différentielle*

La relation $2(2n+3)a_{n+1}=(n+1)a_n$ suggère qu'il peut y avoir une relation simple entre f et f'.

D'après le cours sur les séries entières, la fonction f est de classe $\mathscr{C}^\infty$ sur $]-4\,;4[$ et pour tout $x\in]-4\,;4[$:

$$f(x)=\sum_{n=0}^{+\infty}a_n x^n\quad\text{puis}$$

$$f'(x)=\sum_{n=1}^{+\infty}na_n x^{n-1}=\sum_{n=0}^{+\infty}(n+1)a_{n+1}x^n.$$

On aura donc, pour $x\in]-4\,;4[\setminus\{0\}$:

$$\begin{aligned}4f'(x)&=\sum_{n=0}^{+\infty}4(n+1)a_{n+1}x^n\\&=\sum_{n=0}^{+\infty}2(2n+3)a_{n+1}x^n-2\sum_{n=0}^{+\infty}a_{n+1}x^n\\&=\sum_{n=0}^{+\infty}(n+1)a_n x^n-\frac{2}{x}\sum_{n=0}^{+\infty}a_{n+1}x^{n+1}\\&=\sum_{n=0}^{+\infty}na_n x^n+\sum_{n=0}^{+\infty}a_n x^n-\frac{2}{x}\big(f(x)-a_0\big)\\&=xf'(x)+f(x)-\frac{2f(x)}{x}+\frac{2}{x}\ (a_0=1).\end{aligned}$$

f est donc solution de l'équation différentielle linéaire du 1er ordre :

$$(x^2-4x)y'+(x-2)y+2=0\,,$$

qu'il ne reste plus qu'à résoudre par les méthodes habituelles.

> ✓ Nous vous laissons le soin de résoudre cette équation : d'abord l'équation homogène, puis la méthode de variation de la constante. Il faut simplement faire attention à l'intervalle sur lequel on travaille, à cause des signes et des radicaux qui apparaissent.

Exercice 13.7

• Notons $f\colon t\mapsto\ln t\ln(1-t)$. f est continue sur $]0\,;1[$. De plus :

– $\ln(1-t)\underset{t\to0^+}{\sim}-t$ donc $\lim\limits_{t\to0^+}f(t)=\lim\limits_{t\to0^+}-t\ln t=0$;

– $\ln(t)\underset{t\to1^-}{\sim}t-1$ donc $\lim\limits_{t\to1^-}f(t)=\lim\limits_{t\to1^-}(t-1)\ln(1-t)=0.$

f se prolonge donc en une fonction continue sur $[0\,;1]$ en posant $f(0)=f(1)=0$, ce qui prouve l'existence de l'intégrale I.

• D'après le développement en série entière de $t \mapsto \ln(1-t)$ on a :

$$\forall\, t \in]0\,;1[,\ f(t) = -\sum_{n=1}^{+\infty} \frac{t^n \ln t}{n}\,.$$

Posons pour $n \in \mathbb{N}^*$,

$$u_n(t) = \begin{cases} 0 & \text{si } t = 0 \text{ ou } t = 1 \\ -\dfrac{t^n \ln t}{n} & \text{sinon}\,. \end{cases}$$

Les fonctions u_n sont continues sur $[0\,;1]$ et la série $\sum_{n\geqslant 1} u_n$ converge simplement sur $[0\,;1]$ vers f.
De plus, une rapide étude de fonctions montre que u_n atteint son maximum sur $[0\,;1]$ en $t = \mathrm{e}^{-\frac{1}{n}}$ donc $\|u_n\|_\infty^{[0;1]} = \dfrac{\mathrm{e}^{-1}}{n^2}$.
Ainsi la série est-elle normalement donc uniformément convergente sur $[0\,;1]$. Le théorème rappelé en **[S12.10]** permet d'intervertir $\sum$ et $\int$:

$$I = -\sum_{n=1}^{+\infty} \frac{1}{n} \int_0^1 t^n \ln t \,\mathrm{d}t\,.$$

Une intégration par parties donne alors :

$$\begin{aligned} \int_0^1 t^n \ln t \,\mathrm{d}t &= \left[\frac{t^{n+1}\ln t}{n+1}\right]_0^1 - \int_0^1 \frac{t^n}{n+1}\,\mathrm{d}t \\ &= -\frac{1}{(n+1)^2}\,, \end{aligned}$$

d'où :

$$\begin{aligned} I &= \sum_{n=1}^{+\infty} \frac{1}{n(n+1)^2} \\ &= \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} - \frac{1}{(n+1)^2}\right) \\ &= \underbrace{\sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right)}_{\text{série télescopique}} + 1 - \sum_{n=1}^{+\infty} \frac{1}{n^2} \\ &= 2 - \frac{\pi^2}{6}\,. \end{aligned}$$

✓ Pour ce dernier calcul, on pouvait séparer le terme $\sum_{n=1}^{+\infty} \frac{1}{n^2}$, mais il était impossible de « découper en deux » la somme $\sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right)$!

Exercice 13.8

Pour simplifier, notons, pour tout $n \in \mathbb{N}$, $a_n = \frac{4^n (n!)^2}{(2n+1)!}$, de sorte que la série entière considérée est $\sum_{n\geqslant 0} a_n x^{2n+1}$.

• Pour tout $n \in \mathbb{N}$ et $x \neq 0$, $a_n x^{2n+1}$ ne s'annule pas et l'on a :

$$\frac{\left|a_{n+1}x^{2n+3}\right|}{\left|a_n x^{2n+1}\right|} = \frac{4(n+1)^2}{(2n+2)(2n+3)}\,|x|^2$$

donc $\lim\limits_{n\to+\infty} \dfrac{\left|a_{n+1}x^{2n+3}\right|}{\left|a_n x^{2n+1}\right|} = |x|^2$.

D'après la règle de d'Alembert pour les séries à termes réels positifs :

– si $|x|^2 > 1$, la série $\sum_{n\geqslant 0} \left|a_n x^{2n+1}\right|$ diverge ;

– si $|x|^2 < 1$, elle converge.

Il en résulte que le rayon de convergence de la série considérée est égal à 1.

• En reprenant une partie du calcul précédent, on trouve, pour tout $n \in \mathbb{N}$:

$$\frac{a_{n+1}}{a_n} = \frac{4(n+1)^2}{(2n+2)(2n+3)} = \frac{2(n+1)}{2n+3}\,,$$

ou encore :

$$\forall\, n \in \mathbb{N}\,,\ (2n+3)a_{n+1} = (2n+2)a_n \quad (*)$$

D'après le cours sur les séries entières, si l'on appelle $f(x)$ sa somme, la fonction f est $\mathscr{C}^\infty$ sur $]-1\,;1[$ et, $\forall\, x \in]-1\,;1[$,

$$f(x) = \sum_{n=0}^{+\infty} a_n x^{2n+1} \quad \text{et}$$

$$f'(x) = \sum_{n=0}^{+\infty} (2n+1) a_n x^{2n}\,.$$

Pour tout $x \in]-1\,;1[$, on écrit :

$$\begin{aligned} f'(x) &= a_0 + \sum_{n=1}^{+\infty} (2n+1) a_n x^{2n} \\ &= a_0 + \sum_{n=0}^{+\infty} (2n+3) a_{n+1} x^{2n+2} \\ &= a_0 + \sum_{n=0}^{+\infty} (2n+2) a_n x^{2n+2} \\ &\qquad \text{(grâce à la relation de récurrence } (*)\text{)} \end{aligned}$$

$$f'(x) = a_0 + \sum_{n=0}^{+\infty}(2n+1)a_n x^{2n+2} + \sum_{n=0}^{+\infty} a_n x^{2n+2}$$

(en écrivant $2n+2=(2n+1)+1$)

$$= a_0 + x^2 \underbrace{\sum_{n=0}^{+\infty}(2n+1)a_n x^{2n}}_{=f'(x)} + x \underbrace{\sum_{n=0}^{+\infty} a_n x^{2n+1}}_{=f(x)}$$

Compte tenu de $a_0 = 1$ (calcul direct), on a donc trouvé :

$$f'(x) = 1 + x^2 f'(x) + x f(x)$$

et f est donc la solution sur $]-1\,;1[$ de l'équation différentielle

$$(E) \quad (1-x^2)y' - xy = 1$$

avec $y(0) = a_0 = 1$.

Il ne reste plus qu'à résoudre cette équation. Notons déjà que, puisque le coefficient $(1-x^2)$ de y'' ne s'annule pas sur $]-1\,;1[$, il existe bien une et une seule solution à cette équation qui vérifie la condition initiale $y(0) = 1$, d'après le théorème de Cauchy-Lipschitz.

De plus, cette unique solution sera bien la somme de la série entière cherchée, puisque, en « remontant » les calculs, si une solution de cette équation est développable en série entière sur $]-1\,;1[$, les coefficients a_n doivent vérifier la relation de récurrence $(*)$, et, puisque $a_0 = 1$, ce seront bien les coefficients de départ.

On applique la méthode générale pour résoudre une équation différentielle linéaire du premier ordre, rappelée en **[S17.6]**.

– L'équation homogène associée s'écrit, sur l'intervalle $]-1\,;1[$: $y' = \dfrac{x}{1-x^2}y$.

Ses solutions sont donc les fonctions de la forme :

$$x \mapsto C\exp\left(\int \frac{x}{1-x^2}\,\mathrm{d}x\right) = C\exp\left(-\frac{1}{2}\ln(1-x^2)\right) = \frac{C}{\sqrt{1-x^2}}\,.$$

– Pour résoudre l'équation avec second membre (E), on utilise la méthode de variation de la constante, en cherchant y sous la forme $y(x) = \dfrac{C(x)}{\sqrt{1-x^2}}$, où C est une fonction dérivable sur $]-1\,;1[$ (cela est possible puisque $\sqrt{1-x^2}$ ne s'annule pas sur cet intervalle).

En remplaçant dans l'équation (E), on trouve : $(1-x^2)\dfrac{C'(x)}{\sqrt{1-x^2}} = 1$, soit :

$$C'(x) = \frac{1}{\sqrt{1-x^2}}\,.$$

On en déduit donc :

$$\forall\, x \in\,]-1\,;1[\ ,\ C(x) = \operatorname{Arcsin} x + cste$$

d'où

$$y(x) = \frac{\operatorname{Arcsin} x + cste}{\sqrt{1-x^2}}$$

et la condition $y(0) = 1$ permet alors de conclure : $\forall\, x \in\,]-1\,;1[$,

$$\sum_{n=0}^{+\infty} \frac{4^n (n!)^2}{(2n+1)!} x^{2n+1} = \frac{1+\operatorname{Arcsin} x}{\sqrt{1-x^2}}\,.$$

Exercice 13.9

1) • Soit $a_n = \sin\dfrac{1}{\sqrt{n}}$ pour $n \in \mathbb{N}^*$.

$a_n \underset{n\to+\infty}{\sim} \dfrac{1}{\sqrt{n}}$, donc la série entière $\sum a_n x^n$ a le même rayon de convergence que la série entière $\sum \dfrac{x^n}{\sqrt{n}}$, c'est-à-dire 1.

• L'équivalent ci-dessus montre également que la série diverge pour $x = 1$ (série de Riemann).

• Enfin, pour $x = -1$, la série de terme général $\dfrac{(-1)^n}{\sqrt{n}}$ vérifie le critère spécial des séries alternées (vérification facile), donc converge.

En conclusion, le domaine de définition de f est $[-1\,;1[$.

✓ N'oubliez pas que chercher le domaine de définition ne se limite pas à la recherche du rayon de convergence.

2) La fonction f est croissante sur $[0\,;1[$; cela se vérifie aisément soit en remarquant

que $f'(x) \geqslant 0$, soit directement en vérifiant que :

$$0 \leqslant x \leqslant y < 1 \implies f(x) \leqslant f(y)\,.$$

Si f était majorée sur cet intervalle, il existerait un réel $M \geqslant 0$ tel que :

$$\forall\, x \in [0\,;1[, \sum_{n=1}^{+\infty} \sin\frac{1}{\sqrt{n}}\, x^n \leqslant M$$

donc pour tout $x \in [0\,;1[$ et tout entier $N \in \mathbb{N}^*$ on aurait :

$$\sum_{n=1}^{N} \sin\frac{1}{\sqrt{n}}\, x^n \leqslant M\,.$$

En faisant tendre x vers 1^- dans cette somme finie, on aurait alors :

$$\forall\, N \in \mathbb{N}^*,\ \sum_{n=1}^{N} \sin\frac{1}{\sqrt{n}} \leqslant M\,,$$

et la série de terme général $\sin\frac{1}{\sqrt{n}}$ serait convergente (comme série à termes positifs dont les sommes partielles sont majorées).

On a donc montré par l'absurde que f ne pouvait être majorée sur $[0\,;1[$; étant croissante, le théorème de la limite monotone permet alors d'affirmer que :

$$\lim_{x\to 1^-} f(x) = +\infty\,.$$

3) On a :

$$\begin{aligned}(1-x)f(x) &= \sum_{n=1}^{+\infty} \sin\frac{1}{\sqrt{n}}\, x^n - \sum_{n=1}^{+\infty} \sin\frac{1}{\sqrt{n}}\, x^{n+1}\\ &= (\sin 1)x\\ &\quad + \sum_{n=2}^{+\infty}\left(\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}}\right)x^n.\end{aligned}$$

D'après l'inégalité des accroissements finis, la dérivée de sin étant majorée par 1 en valeur absolue, on a pour $n \geqslant 2$:

$$\begin{aligned}\left|\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}}\right| &\leqslant \frac{1}{\sqrt{n-1}} - \frac{1}{\sqrt{n}}\\ &\leqslant \frac{\sqrt{n}-\sqrt{n-1}}{\sqrt{n}\sqrt{n-1}}\\ &\leqslant \frac{1}{\left(\sqrt{n}+\sqrt{n-1}\right)\sqrt{n}\sqrt{n-1}}\,,\end{aligned}$$

ce qui montre que :

$$\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}} = O\left(\frac{1}{n^{3/2}}\right)$$

(ce résultat peut aussi s'obtenir avec un développement limité).

La série :

$$\sum_{n=2}^{+\infty}\left(\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}}\right)x^n$$

est donc normalement convergente sur $[0\,;1]$. Elle est donc aussi uniformément convergente, donc le théorème de la double limite s'applique et donne

$$\begin{aligned}\lim_{x\to 1}\left(\sum_{n=2}^{+\infty}\left(\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}}\right)x^n\right)&\\ = \sum_{n=2}^{+\infty}\left(\sin\frac{1}{\sqrt{n}} - \sin\frac{1}{\sqrt{n-1}}\right)&\\ = -\sin 1 \quad \text{(télescopage)}.&\end{aligned}$$

Compte tenu du calcul fait au début de la question, on en déduit :

$$\lim_{x\to 1}(1-x)f(x) = 0\,.$$

Exercice 13.10

La relation de l'énoncé peut s'écrire :

$$\forall\, n \in \mathbb{N},\ 2\frac{a_{n+1}}{n!} = \sum_{k=0}^{n} \frac{a_k}{k!}\,\frac{a_{n-k}}{(n-k)!}\,.$$

Cela fait penser à considérer la série entière $\sum_{n\geqslant 0} \frac{a_n}{n!}x^n$. En supposant son rayon de convergence $R > 0$, notons :

$$\forall\, x \in\,]-R\,;R[,\ f(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n!}x^n\,.$$

La série entière $\sum_{n\geqslant 0} c_n x^n$ avec :

$$\forall\, n \in \mathbb{N},\ c_n = \sum_{k=0}^{n} \frac{a_k}{k!}\,\frac{a_{n-k}}{(n-k)!}$$

n'est autre que la série produit de Cauchy de $\sum_{n\geqslant 0} \frac{a_n}{n!}x^n$ par elle-même. D'après le cours, elle a donc un rayon de convergence supérieur ou égal à R et :

$$\forall\, x \in\,]-R\,;R[,\ \sum_{n=0}^{+\infty} c_n x^n = f(x)^2\,.$$

D'autre part, par le théorème de dérivation d'une série entière on obtient, pour $x \in\,]-R\,;R[$:

$$f'(x) = \sum_{n=1}^{+\infty} \frac{a_n}{(n-1)!}x^{n-1} = \sum_{n=0}^{+\infty} \frac{a_{n+1}}{n!}x^n,$$

donc la relation de l'énoncé implique :

$$\forall\, x \in]-R\,;R[,\ 2f'(x) = f(x)^2\,.$$

Puisque $f(0) = 1$, il existe un voisinage V de 0 sur lequel f ne s'annule pas. L'équation précédente s'écrit alors :

$$\forall\, x \in V,\ -\frac{f'(x)}{f(x)^2} = -\frac{1}{2}$$

donc il existe une constante C telle que

$$\forall\, x \in V,\ \frac{1}{f(x)} = -\frac{x}{2} + C$$

et compte tenu de $f(0) = a_0 = 1$ on trouve :

$$f(x) = \frac{1}{1 - \frac{x}{2}}\,\cdot$$

Réciproquement, la fonction

$f\colon x \mapsto \dfrac{1}{1-\frac{x}{2}}$ admet le développement en série entière :

$$\forall\, x \in]-2\,;2[,\ f(x) = \sum_{n=0}^{+\infty} \left(\frac{x}{2}\right)^n\,.$$

On doit donc avoir :

$$\forall\, n \in \mathbb{N},\ a_n = \frac{n!}{2^n}\,,$$

et il est facile de vérifier directement dans la relation de récurrence proposée que cette suite convient.

Exercice 14.1

1) On commence *déjà* par remarquer que f est définie et continue sur $]1\,;+\infty[$ (en effet, $\ln x$ est bien positif sur cet intervalle).

> ✓ Rappelons que la *première* chose à faire lors de l'étude d'une intégrale généralisée est d'étudier le domaine de définition et la continuité de la fonction considérée.

f étant continue sur $]1\,;+\infty[$, les problèmes d'intégrabilité ne se posent donc qu'aux voisinages de 1^+ et de $+\infty$. De plus, f étant à valeurs positives, étudier son intégrabilité équivaut à démontrer simplement la convergence de l'intégrale.

• On commence, lors de cette étude, par examiner la fonction et se « débarrasser » des parties qui ne posent pas de problème. Ici, $\lim\limits_{x\to 1} \sqrt{x} = 1$ donc on a $f(x) \underset{x\to 1^+}{\sim} \dfrac{\sqrt{\ln x}}{x-1}$.

Ensuite, on dispose de l'équivalent : $\ln x \underset{x\to 1^+}{\sim} x-1$, donc $\sqrt{\ln x} \underset{x\to 1^+}{\sim} \sqrt{x-1}$ et $f(x) \underset{x\to 1^+}{\sim} \dfrac{1}{\sqrt{x-1}}\,\cdot$

Or, la fonction $x \mapsto \dfrac{1}{\sqrt{x-1}} = \dfrac{1}{(x-1)^{1/2}}$ est intégrable au voisinage de 1^+ : c'est une intégrale de Riemann, de même nature que celle de $x \mapsto \dfrac{1}{x^{1/2}}$ au voisinage de 0^+, donc convergente ($\frac{1}{2} < 1$).

Puisqu'il s'agit d'une fonction positive (critère à ne pas oublier !), les théorèmes de comparaison assurent que f est aussi intégrable au voisinage de 1^+.

• Au voisinage de $+\infty$, on a $x-1 \sim x$, donc $f(x) \underset{x\to+\infty}{\sim} \dfrac{\sqrt{\ln x}}{x^{3/2}}$. Il s'agit là d'une intégrale de Bertrand, dont l'étude est classique (nous avons déjà étudié ces intégrales, vous reporter en **[SF14.3]** page 539).

L'étude de l'intégrabilité de cette fonction consiste à la comparer en $+\infty$ à une fonction de la forme $x \mapsto \dfrac{1}{x^\alpha}$ avec $\alpha > 1$. Or $x^\alpha f(x) \underset{+\infty}{\sim} \dfrac{\sqrt{\ln x}}{x^{3/2-\alpha}}$, donc si l'on choisit α quelconque dans $\left]1\,;\frac{3}{2}\right[$, on aura, d'après les croissances comparées, $\lim\limits_{x\to+\infty} x^\alpha f(x) = 0$.

Donc $f(x) \underset{x\to+\infty}{=} o\left(\dfrac{1}{x^\alpha}\right)$ avec $\alpha > 1$; or $x \mapsto \dfrac{1}{x^\alpha}$ est positive et intégrable au voisinage de $+\infty$, donc il en est de même de f.

En rassemblant les deux cas, on conclut : f est intégrable sur $]1\,;+\infty[$.

2) La fonction g est définie et continue sur $]0\,;1[$ puisque $|\ln x|$ est bien positif sur cet intervalle.
De plus, g est de signe constant sur $]0\,;1[$ donc l'étude de son intégrabilité équivaut à démontrer la convergence de l'intégrale.

• L'intégrabilité de g au voisinage de 1 se démontre comme celle de f ci-dessus.

• Au voisinage de 0, on a :

$$g(x) \underset{x\to 0^+}{\sim} -\frac{\sqrt{|\ln x|}}{\sqrt{x}}\,,$$

et il s'agit là encore d'une intégrale de Bertrand. On applique donc le même principe qu'auparavant.
On remarque ici que :

$$\lim_{x\to 0^+} x^{3/4} g(x) = \lim_{x\to 0^+} x^{1/4}\sqrt{|\ln x|} = 0\,,$$

donc $g(x) \underset{x\to 0^+}{=} o\left(\dfrac{1}{x^{3/4}}\right)$.
Puisque la fonction $x \mapsto \dfrac{1}{x^{3/4}}$ est intégrable au voisinage de 0^+, il en est de même de g.
En conclusion, g est intégrable sur $]0\,;1[$.

Exercice 14.2

• La fonction $t \mapsto \operatorname{th} t$ est une application continue de $\mathbb{R}_+^*$ dans $]0\,;1[$, donc l'application $t \mapsto \ln(\operatorname{th} t)$ est bien continue sur $\mathbb{R}_+^*$.

• La fonction $t \mapsto \operatorname{th} t$ étant de plus une bijection strictement croissante et de classe $\mathscr{C}^1$ de $\mathbb{R}_+^*$ sur $]0\,;1[$, le théorème **[S14.5]** permet directement d'affirmer que l'intégrale proposée est de même nature que celle obtenue par le changement de variable $u = \operatorname{th} t$.

Or ce changement de variable conduit à :

$$\int_0^{+\infty} \ln(\operatorname{th} t)\,\mathrm{d}t \;= \int_0^1 \frac{\ln u}{1-u^2}\,\mathrm{d}u\,,$$

puisque la dérivée de la fonction th est la fonction $t \mapsto 1 - \operatorname{th}^2 t$.

• On est donc ramené à étudier l'existence de l'intégrale :

$$\int_0^1 \frac{\ln u}{1-u^2}\,\mathrm{d}u\,.$$

La fonction à intégrer étant continue et de signe constant sur $]0\,;1[$, il suffit alors de remarquer que :

– au voisinage de 0^+, $\dfrac{\ln u}{1-u^2} \underset{u\to 0^+}{\sim} \ln u$, et puisque la fonction ln est intégrable au voisinage de 0^+, il en est de même de $u \mapsto \dfrac{\ln u}{1-u^2}$;

– au voisinage de 1^-,

$$\lim_{u\to 1^-} \frac{\ln u}{1-u^2} = \lim_{u\to 1^-} \frac{-\ln u}{u-1}\cdot\frac{1}{1+u} = -\frac{1}{2},$$

donc la fonction se prolonge par continuité en 1^-, et par suite est intégrable au voisinage de 1.

• En conclusion, l'intégrale proposée existe.

✓ On pourrait même calculer l'intégrale précédente, en utilisant le développement en série entière de $\dfrac{1}{1-u^2}$ et une interversion série-intégrale.

Exercice 14.3

La fonction $t \mapsto te^{-t^3|\sin t|}$ est continue sur $\mathbb{R}_+$. Pour tout $x \geqslant 0$ posons :

$$F(x) = \int_0^x te^{-t^3|\sin t|}\,\mathrm{d}t\,.$$

Il s'agit donc d'étudier $\lim\limits_{x\to+\infty} F(x)$.

• Soit $n \in \mathbb{N}^*$. Par la relation de Chasles, $F(n\pi) = \sum\limits_{k=0}^{n-1} u_k$ où

$$u_k = \int_{k\pi}^{(k+1)\pi} te^{-t^3|\sin t|}\,\mathrm{d}t\,.$$

En faisant le changement de variable $u = t - k\pi$, on obtient :

$$\begin{aligned} u_k &= \int_0^\pi (u+k\pi)e^{-(u+k\pi)^3|\sin(u+k\pi)|}\,\mathrm{d}u \\ &= \int_0^\pi (u+k\pi)e^{-(u+k\pi)^3\sin u}\,\mathrm{d}u\,. \end{aligned}$$

Puisque $k\pi \leqslant u + k\pi \leqslant (k+1)\pi$ lorsque $u \in [0\,;\pi]$, on a l'encadrement :

$$0 \leqslant u_k \leqslant (k+1)\pi \int_0^\pi e^{-k^3\pi^3\sin u}\,\mathrm{d}u\,.$$

Or, par le changement de variable $u \mapsto \pi - u$, on a :

$$\int_0^{\frac{\pi}{2}} e^{-k^3\pi^3\sin u}\,\mathrm{d}u = \int_{\frac{\pi}{2}}^\pi e^{-k^3\pi^3\sin u}\,\mathrm{d}u\,,$$

donc

$$0 \leqslant u_k \leqslant 2(k+1)\pi \int_0^{\frac{\pi}{2}} e^{-k^3\pi^3\sin u}\,\mathrm{d}u\,.$$

Or pour $u \in \left[0\,;\frac{\pi}{2}\right]$, $\sin u \geqslant \dfrac{2}{\pi}u$ (car la fonction sin est concave sur cet intervalle, donc située au-dessus de ses cordes), d'où :

$$0 \leqslant u_k \leqslant 2(k+1)\pi \int_0^{\frac{\pi}{2}} e^{-2k^3\pi^2 u}\,\mathrm{d}u\,.$$

Le calcul de cette intégrale, qui est simple et que nous ne reproduisons pas ici, permet alors d'en déduire :

$$u_k \underset{k\to+\infty}{=} O\left(\frac{1}{k^2}\right).$$

D'après les théorèmes des comparaison des séries à termes positifs, la série $\sum u_k$ est convergente, ce qui signifie que $\lim\limits_{n\to+\infty} F(n\pi)$ existe.

• Notons $\ell = \lim\limits_{n\to+\infty} F(n\pi)$. Pour tout $x \geqslant 0$, il existe un et un seul entier n_x tel que $n_x\pi \leqslant x < (n_x+1)\pi$; plus précisément, $n_x = \left\lfloor \dfrac{x}{\pi} \right\rfloor$.

La fonction F étant croissante (car c'est une primitive d'une fonction positive), on a l'encadrement :

$$F(n_x\pi) \leqslant F(x) \leqslant F((n_x+1)\pi)$$

et puisque n_x tend vers $+\infty$ avec x, on en déduit $\lim\limits_{x\to+\infty} F(x) = \ell$.

En conclusion, l'intégrale proposée existe.

Exercice 14.4

1) Pour $x \geqslant 0$, soit $F(x) = \displaystyle\int_0^x f$. F est une primitive de f continue donc est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$.

Puisque $g(x) = \dfrac{1}{x}F(x)$ pour $x > 0$, g est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$. Puis :

$$\begin{aligned}\lim_{x\to0^+} g(x) &= \lim_{x\to0^+} \frac{F(x)-F(0)}{x-0}\\ &= F'(0) = f(0) = g(0)\,,\end{aligned}$$

ce qui prouve que g est aussi continue en 0, donc sur $\mathbb{R}_+$.

2) a) Il faut évidemment penser à faire une intégration par parties. En utilisant le fait qu'une primitive de f est $t \mapsto tg(t)$ et que, pour $x > 0$:

$$\begin{aligned}g'(x) &= \frac{f(x)}{x} - \frac{1}{x^2}\int_0^x f(t)\,\mathrm{d}t\\ &= \frac{f(x)}{x} - \frac{g(x)}{x}\,,\end{aligned}$$

on obtient, pour $0 < a < b$:

$$\begin{aligned}&\int_a^b \underbrace{f(t)}_{u'(t)}\underbrace{g(t)}_{v(t)}\,\mathrm{d}t\\ &= \left[tg^2(t)\right]_a^b - \int_a^b \underbrace{tg(t)}_{u(t)}\underbrace{\left(\frac{f(t)}{t} - \frac{g(t)}{t}\right)}_{v'(t)}\,\mathrm{d}t\\ &= bg^2(b) - ag^2(a)\\ &\qquad - \int_a^b f(t)g(t)\,\mathrm{d}t + \int_a^b g^2(t)\,\mathrm{d}t\end{aligned}$$

ce qui donne l'égalité de l'énoncé.

b) Les fonctions f et g étant continues sur $\mathbb{R}_+$, on peut faire tendre a vers 0 dans l'égalité précédente, ce qui donne :

$$\int_0^b g^2(t)\,\mathrm{d}t = 2\int_0^b f(t)g(t)\,\mathrm{d}t - bg^2(b)\,.$$

On en déduit que :

$$\int_0^b g^2(t)\,\mathrm{d}t \leqslant 2\int_0^b f(t)g(t)\,\mathrm{d}t\,,$$

puis, en utilisant l'inégalité de Cauchy-Schwarz :

$$\int_0^b g^2 \leqslant 2\sqrt{\int_0^b f^2}\sqrt{\int_0^b g^2}.$$

On peut bien sûr supposer g non identiquement nulle (cela donnerait $f = 0$, et l'exercice n'a alors pas d'intérêt), donc (théorème classique) l'intégrale $\displaystyle\int_0^b g^2(t)\,\mathrm{d}t$ n'est pas nulle, et l'inégalité précédente se simplifie en :

$$\sqrt{\int_0^b g^2(t)\,\mathrm{d}t} \leqslant 2\sqrt{\int_0^b f^2(t)\,\mathrm{d}t},$$

d'où enfin :

$$\int_0^b g^2(t)\,\mathrm{d}t \leqslant 4\int_0^b f^2(t)\,\mathrm{d}t \leqslant 4\int_0^{+\infty} f^2(t)\,\mathrm{d}t$$

puisque f^2 est intégrable sur $\mathbb{R}_+$.

Les intégrales $\displaystyle\int_0^b g^2(t)\,\mathrm{d}t$ de la fonction continue positive g^2 étant majorées pour tout $b \in \mathbb{R}_+$, cela implique que g^2 est intégrable sur $\mathbb{R}_+$, et on a en plus l'inégalité :

$$\int_0^{+\infty} g^2(t)\,\mathrm{d}t \leqslant 4\int_0^{+\infty} f^2(t)\,\mathrm{d}t\,.$$

(*inégalités de Hardy*).

c) • On sait que f^2 est intégrable sur $\mathbb{R}_+$, et on vient de démontrer que g^2 l'est également.

Le théorème de comparaison pour les intégrales de fonctions positives et l'inégalité :

$$|f(t)g(t)| \leqslant \frac{1}{2}\left(|f(t)|^2 + |g(t)|^2\right).$$

permettent d'en déduire que fg est intégrable sur $\mathbb{R}_+$.

• On rappelle la relation vue plus haut :

$$\int_0^b g^2(t)\,\mathrm{d}t = 2\int_0^b f(t)g(t)\,\mathrm{d}t - bg^2(b)\,.$$

Puisque, d'après ce qui vient d'être démontré, les deux intégrales dans cette égalité ont une limite quand b tend vers $+\infty$, il en résulte que $\lim\limits_{b\to+\infty} bg^2(b)$ existe.
Notons ℓ cette limite.
Si ℓ était non nulle, on pourrait écrire $g^2(t) \underset{t\to+\infty}{\sim} \frac{\ell}{t}$; cela est impossible puisque g^2 est intégrable sur $\mathbb{R}_+$ alors que la fonction $t \mapsto \frac{1}{t}$ n'est pas intégrable au voisinage de $+\infty$. On a donc forcément $\ell = 0$, soit :

$$\lim_{x\to+\infty} xg^2(x) = 0\,.$$

Exercice 14.5

• On notera, pour tout $x \geqslant 1$:

$$F(x) = \int_1^x f(t)\,\mathrm{d}t\,.$$

Supposons d'abord que $\lim\limits_{x\to+\infty} \dfrac{F(x)}{x} = \ell$ existe. Par une intégration par parties :

$$\int_x^y \frac{f(t)}{t^2}\,\mathrm{d}t = \left[\frac{F(t)}{t^2}\right]_x^y + 2\int_x^y \frac{F(t)}{t^3}\,\mathrm{d}t\,.$$

Puisque $F(t) \underset{+\infty}{=} O(t)$, $\dfrac{F(t)}{t^3} = O\left(\dfrac{1}{t^2}\right)$ et par les théorèmes de comparaison, l'intégrale $\int_x^{+\infty} \frac{F(t)}{t^3}$ existe.
De plus, $\lim\limits_{y\to+\infty} \dfrac{F(y)}{y^2} = 0$ donc $\int_x^{+\infty} \dfrac{f(t)}{t^2}\,\mathrm{d}t$ existe, et en faisant $y \to +\infty$ dans la relation précédente on obtient :

$$x\int_x^{+\infty}\frac{f(t)}{t^2}\,\mathrm{d}t = -\frac{F(x)}{x} + 2x\int_x^{+\infty}\frac{F(t)}{t^3}\,\mathrm{d}t\,.$$

La propriété $\lim\limits_{t\to+\infty} \dfrac{F(t)}{t} = \ell$ peut aussi s'écrire :

$$\frac{F(t)}{t} - \ell \underset{+\infty}{=} o(1)\ ,$$

d'où

$$\frac{F(t)}{t^3} - \frac{\ell}{t^2} \underset{+\infty}{=} o\left(\frac{1}{t^2}\right).$$

L'intégration des relations de comparaison pour les intégrales convergentes permet alors d'écrire :

$$\int_x^{+\infty}\left(\frac{F(t)}{t^3} - \frac{\ell}{t^2}\right)\mathrm{d}t \underset{+\infty}{=} o\left(\int_x^{+\infty}\frac{\mathrm{d}t}{t^2}\right),$$

d'où :

$$\int_x^{+\infty}\frac{F(t)}{t^3}\,\mathrm{d}t - \frac{\ell}{x} \underset{x\to+\infty}{=} o\left(\frac{1}{x}\right).$$

On en déduit :

$$\lim_{x\to+\infty} 2x\int_x^{+\infty}\frac{F(t)}{t^3}\,\mathrm{d}t = 2\ell$$

puis :

$$\lim_{x\to+\infty} x\int_x^{+\infty}\frac{f(t)}{t^2}\,\mathrm{d}t = \ell\,.$$

✓ Il aurait été plus rapide d'utiliser directement la relation $F(x) \underset{+\infty}{\sim} \ell x$ pour l'intégration des relations de comparaison, mais cette écriture n'a pas de sens lorsque $\ell = 0$. C'est pour cette raison que nous avons traduit la limite à l'aide de « o ».

• Supposons maintenant que :

$$\lim_{x\to+\infty} x\int_x^{+\infty}\frac{f(t)}{t^2}\,\mathrm{d}t = \ell$$

existe.
Notons pour $x \geqslant 1$, $G(x) = \int_x^{+\infty} \frac{f(t)}{t^2}\,\mathrm{d}t$; ainsi G est de classe $\mathscr{C}^1$ et $G'(x) = -\dfrac{f(x)}{x^2}$, soit $f(x) = -x^2G'(x)$.
En intégrant par parties :

$$\begin{aligned}\int_1^x f(t)\,\mathrm{d}t &= \left[-t^2G(t)\right]_1^x + 2\int_1^x tG(t)\,\mathrm{d}t\\ &= G(1) - x^2G(x) + 2\int_1^x tG(t)\,\mathrm{d}t\ (*)\end{aligned}$$

Par hypothèse, $\lim\limits_{t\to+\infty} tG(t) = \ell$, ce qui s'écrit $tG(t) - \ell \underset{+\infty}{=} o(1)$ et par intégration des relations de comparaison dans le cas d'intégrales divergentes, on obtient :

$$\int_1^x tG(t)\,\mathrm{d}t - \int_1^x \ell\,\mathrm{d}t \underset{+\infty}{=} o\left(\int_1^x \mathrm{d}t\right),$$

soit :

$$\int_1^x tG(t)\,\mathrm{d}t - \ell(x-1) \underset{+\infty}{=} o(x)\ .$$

On en déduit $\lim\limits_{x\to+\infty} \dfrac{1}{x}\int_1^x tG(t)\,\mathrm{d}t = \ell$ puis la relation $(*)$ conduit (en la divisant par x) à :

$$\lim_{x\to+\infty}\frac{1}{x}\int_0^x f(t)\,\mathrm{d}t = -\ell + 2\ell = \ell\,,$$

ce qu'il fallait démontrer.

Exercice 14.6

Soit $F : x \mapsto \int_1^x f(t)\,\mathrm{d}t$ pour $x \geqslant 1$. F est de classe $\mathscr{C}^1$ sur $[1\,;+\infty]$ et en intégrant par parties :

$$\int_1^x \frac{f(t)}{t^\alpha}\,\mathrm{d}t = \frac{F(x)}{x^\alpha} + \alpha\int_1^x \frac{F(t)}{t^{\alpha+1}}\,\mathrm{d}t\,.$$

Puisque l'intégrale de f converge, $\lim\limits_{x\to+\infty} F(x)$ existe donc :

$$\lim_{x\to+\infty} \frac{F(x)}{x^\alpha} = 0\,.$$

On a aussi F bornée sur $[1\,;+\infty]$ (car continue et possède une limite en $+\infty$). On a donc :

$$\forall t \geqslant 1,\ \left|\frac{F(t)}{t^{\alpha+1}}\right| \leqslant \frac{\|F\|_\infty}{t^{\alpha+1}}\,.$$

Par comparaison à une fonction intégrable positive ($\alpha+1>1$), on en déduit que l'intégrale $\int_1^{+\infty} \frac{F(t)}{t^{\alpha+1}}\,\mathrm{d}t$ est absolument convergente.

En conclusion, $\int_1^{+\infty} \frac{f(t)}{t^\alpha}\,\mathrm{d}t$ existe, et :

$$\int_1^{+\infty} \frac{f(t)}{t^\alpha}\,\mathrm{d}t = \alpha \int_1^{+\infty} \frac{F(t)}{t^{\alpha+1}}\,\mathrm{d}t\,.$$

Exercice 14.7

1) L'hypothèse de l'énoncé peut s'écrire :

$$\frac{xf'(x)}{f(x)} \underset{x\to+\infty}{=} \alpha + o(1)\ ,$$

ou encore :

$$\frac{f'(x)}{f(x)} - \frac{\alpha}{x} \underset{x\to+\infty}{=} o\left(\frac{1}{x}\right).$$

La fonction $x \mapsto \frac{1}{x}$ étant positive et non intégrable au voisinage de $+\infty$, le théorème d'intégration des relations de comparaison donne :

$$\int_1^x \left(\frac{f'(t)}{f(t)} - \frac{\alpha}{t}\right)\mathrm{d}t \underset{x\to+\infty}{=} o\left(\int_1^x \frac{\mathrm{d}t}{t}\right),$$

d'où :

$$\ln\bigl(f(x)\bigr) - \alpha\ln x \underset{x\to+\infty}{=} o(\ln x)$$

(les constantes d'intégration sont en effet négligeables devant $\ln x$).
On en déduit :

$$\lim_{x\to+\infty} \frac{\ln f(x)}{\ln x} = \alpha\,.$$

2) • Pour montrer l'intégrabilité de f, on la compare à une fonction de Riemann.
Si $\alpha<-1$, il existe β tel que $\alpha<\beta<-1$. On a alors :

$$\begin{aligned}\ln\bigl(x^{-\beta}f(x)\bigr) &= \ln\bigl(f(x)\bigr) - \beta\ln x\\ &= (\alpha-\beta)\ln x + o(\ln x)\ ,\end{aligned}$$

donc $\lim\limits_{x\to+\infty} \ln\bigl(x^{-\beta}f(x)\bigr) = -\infty$, soit $\lim\limits_{x\to+\infty} x^{-\beta}f(x) = 0$ et $f(x) \underset{x\to+\infty}{=} o\Bigl(\frac{1}{x^{-\beta}}\Bigr)$.
Puisque $-\beta>1$, la fonction $x \mapsto \frac{1}{x^{-\beta}}$ est intégrable au voisinage de $+\infty$, donc f aussi.

> ✓ Remarquez que la méthode utilisée est tout à fait similaire à celle utilisée pour les intégrales de Bertrand.

• Puisque $\alpha\neq 0$, l'hypothèse de l'énoncé s'écrit aussi :

$$xf'(x) \underset{x\to+\infty}{\sim} \alpha f(x)\,.$$

Puisque f est positive et intégrable sur $\mathbb{R}_+$, il en est de même de $x \mapsto xf'(x)$ et le théorème d'intégration des relations de comparaison donne :

$$\int_x^{+\infty} tf'(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} \alpha\int_x^{+\infty} f(t)\,\mathrm{d}t\,.$$

Enfin, puisque $f(x) \underset{x\to+\infty}{=} o\Bigl(\frac{1}{x^{-\beta}}\Bigr)$ avec $-\beta>1$, on a $\lim\limits_{x\to+\infty} xf(x) = 0$ donc une intégration par parties donne :

$$\begin{aligned}\int_x^{+\infty} tf'(t)\,\mathrm{d}t &= \bigl[tf(t)\bigr]_x^{+\infty} - \int_x^{+\infty} f(t)\,\mathrm{d}t\\ &= -xf(x) - \int_x^{+\infty} f(t)\,\mathrm{d}t.\end{aligned}$$

En rassemblant les deux résultats précédents, on obtient :

$$\alpha\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{+\infty}{\sim} -xf(x) - \int_x^{+\infty} f(t)\,\mathrm{d}t\,,$$

donc puisque $\alpha+1\neq 0$:

$$(\alpha+1)\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} -xf(x)\,.$$

3) Le principe est le même ici que dans la question précédente, nous détaillerons donc un peu moins.
• Soit β tel que $-1<\beta<\alpha$. Alors on trouve ici :

$$\lim_{x\to+\infty} \ln\bigl(x^{-\beta}f(x)\bigr) = +\infty,$$

donc $\frac{1}{x^{-\beta}} \underset{+\infty}{=} o(f(x))$.
Puisque $-\beta<1$, on en déduit que f n'est pas intégrable au voisinage de $+\infty$.

• L'hypothèse de l'énoncé peut s'écrire :

$$xf'(x) - \alpha f(x) \underset{x\to+\infty}{=} o\big(f(x)\big)\,,$$

et comme f est positive et non intégrable su $\mathbb{R}_+$, le théorème d'intégration des relations de comparaisons donne ici :

$$\int_0^x \big(tf'(t) - \alpha f(t)\big)\,\mathrm{d}t \underset{+\infty}{=} o\bigg(\int_0^x f(t)\,\mathrm{d}t\bigg).$$

Or en intégrant par parties :

$$\int_0^x tf'(t)\,\mathrm{d}t = xf(x) - \int_0^x f(t)\,\mathrm{d}t$$

donc

$$xf(x) - (\alpha+1)\int_0^x f(t)\,\mathrm{d}t \underset{+\infty}{=} o\bigg(\int_0^x f(t)\,\mathrm{d}t\bigg),$$

ce qui signifie que :

$$(\alpha+1)\int_0^x f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} xf(x)\,,$$

ce qu'il fallait démontrer.

Exercice 14.8

1) Nous allons ici scrupuleusement vérifier les hypothèses du théorème de dérivabilité d'une intégrale à paramètre.
Notons, pour $(x,t) \in \mathbb{R}\times\mathbb{R}$, $g(x,t) = \mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-t^2}$.

• Pour tout $x \in \mathbb{R}$, la fonction $t \mapsto g(x,t)$ est continue sur $\mathbb{R}$, et $|g(x,t)| \leqslant \mathrm{e}^{-t^2}$. Or, la fonction $t \mapsto \mathrm{e}^{-t^2}$ est intégrable sur $\mathbb{R}$: en effet, elle y est continue, donc le problème d'intégrabilité ne se pose qu'aux voisinages de $\pm\infty$, et $\mathrm{e}^{-t^2} \underset{\pm\infty}{=} o\Big(\frac{1}{t^2}\Big)$ avec $t \mapsto \frac{1}{t^2}$ positive et intégrable au voisinage de $\pm\infty$.

Il résulte des théorèmes de comparaison pour les intégrales impropres que $t \mapsto g(x,t)$ est intégrable sur $\mathbb{R}$.

• Pour tout $t \in \mathbb{R}$, la fonction g admet une dérivée partielle par rapport à x : $\dfrac{\partial g}{\partial x}(x,t) = \mathrm{i}t\mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-t^2}$.
De façon immédiate, les deux hypothèses suivantes sont vérifiées :

– pour tout $x \in \mathbb{R}$, la fonction $t \mapsto \dfrac{\partial g}{\partial x}(x,t)$ est continue sur $\mathbb{R}$;

– pour tout $t \in \mathbb{R}$, la fonction $x \mapsto \dfrac{\partial g}{\partial x}(x,t)$ est continue sur $\mathbb{R}$.

• Il reste l'hypothèse de domination.
On a : $\forall\,(x,t) \in \mathbb{R}^2$, $\left|\dfrac{\partial g}{\partial x}(x,t)\right| \leqslant t\mathrm{e}^{-t^2}$, et la fonction $t \mapsto t\mathrm{e}^{-t^2}$ est intégrable sur $\mathbb{R}$ (la justification est la même que pour $t \mapsto \mathrm{e}^{-t^2}$ ci-dessus).

Le théorème du cours permet alors de conclure :

$$f \text{ est de classe } \mathscr{C}^1 \text{ sur } \mathbb{R}, \text{ et } \forall\, x \in \mathbb{R},\ f'(x) = \int_{-\infty}^{+\infty} \mathrm{i}t\mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-t^2}\,\mathrm{d}t.$$

2) a) Pour trouver une relation entre f et f', nous allons utiliser une intégration par parties(en la justifiant). On a :

$$\int_{-\infty}^{+\infty} \underbrace{t\mathrm{e}^{-t^2}}_{u'(t)}\underbrace{\mathrm{i}\mathrm{e}^{\mathrm{i}tx}}_{v(t)}\mathrm{d}t = \left[\frac{-1}{2}\mathrm{e}^{-t^2}\mathrm{i}\mathrm{e}^{\mathrm{i}tx}\right]_{-\infty}^{+\infty} - \frac{x}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t\,,$$

et cette écriture a bien un sens puisque $\lim\limits_{t\to\pm\infty} \mathrm{e}^{-t^2}\mathrm{e}^{\mathrm{i}tx} = 0$ (car la fonction $t \mapsto \mathrm{e}^{\mathrm{i}tx}$ est bornée). On a donc :

$$\int_{-\infty}^{+\infty} t\mathrm{e}^{-t^2}\mathrm{i}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t = -\frac{x}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\mathrm{e}^{\mathrm{i}tx}\,\mathrm{d}t$$

d'où :

$$f'(x) = -\frac{x}{2}f(x)\,.$$

b) La solution générale de cette équation différentielle linéaire du premier ordre est $f(x) = C\mathrm{e}^{-x^2/4}$, où C est une constante.
Puisque l'énoncé donne $f(0) = \sqrt{\pi}$, on trouve :

$$\forall\, x \in \mathbb{R}\ ,\ f(x) = \sqrt{\pi}\mathrm{e}^{-x^2/4}\,.$$

3) • Pour tout $x > 0$, la fonction $t \mapsto \mathrm{e}^{-xt}\sin\sqrt{t}$ est continue sur $[0\,;+\infty[$; de plus, $\left|\mathrm{e}^{-xt}\sin\sqrt{t}\right| \leqslant \mathrm{e}^{-xt}$, et, puisque $t \mapsto \mathrm{e}^{-xt}$ est intégrable sur $\mathbb{R}_+$ (car $x > 0$), $g(x)$ est bien définie pour tout $x > 0$.

• En effectuant dans l'intégrale le changement de variable $t = u^2$, qui est une bijection de classe $\mathscr{C}^1$ strictement monotone de $\mathbb{R}_+^*$ sur $\mathbb{R}_+^*$, on obtient :

$$\forall\, x > 0,\ g(x) = \int_0^{+\infty} \mathrm{e}^{-xt}\sin\sqrt{t}\,\mathrm{d}t = \int_0^{+\infty} 2u\mathrm{e}^{-xu^2}\sin u\,\mathrm{d}u\,.$$

On effectue alors une intégration par parties :

$$\begin{aligned}\int_0^{+\infty} 2ue^{-xu^2}\sin u\,du &= \left[-\frac{1}{x}e^{-xu^2}\sin u\right]_0^{+\infty} \\ &\quad + \frac{1}{x}\int_0^{+\infty} e^{-xu^2}\cos u\,du \\ &= 0 + \frac{1}{x}\int_0^{+\infty} e^{-xu^2}\cos u\,du\end{aligned}$$

car la fonction sin étant bornée, on a $\lim\limits_{u\to+\infty} e^{-xu^2}\sin u = 0$ (car $x>0$). Ainsi :

$$\begin{aligned}g(x) &= \frac{1}{x}\int_0^{+\infty} e^{-xu^2}\cos u\,du \\ &= \frac{1}{x^{3/2}}\int_0^{+\infty} e^{-t^2}\cos\left(\frac{t}{\sqrt{x}}\right)dt \quad (1)\end{aligned}$$

en ayant fait encore un changement de variable : $u = \dfrac{t}{\sqrt{x}}$.

Or :

$$\begin{aligned}f(x) &= \int_{-\infty}^{+\infty} e^{itx}e^{-t^2}\,dt \\ &= \int_{-\infty}^{+\infty} e^{-t^2}\cos(tx)\,dt \\ &\quad + i\int_{-\infty}^{+\infty} e^{-t^2}\sin(tx)\,dt \\ &= 2\int_0^{+\infty} e^{-t^2}\cos(tx)\,dt\end{aligned}$$

puisque la première de ces intégrales est celle d'une fonction paire et que la seconde est celle d'une fonction impaire intégrables sur $\mathbb{R}$.

En comparant avec (1), on obtient : $g(x) = \frac{1}{2x^{3/2}} f\left(\frac{1}{\sqrt{x}}\right)$.

Finalement :

$$\forall\, x>0,\ g(x) = \frac{\sqrt{\pi}}{2x^{3/2}}e^{-\frac{1}{4x}} .$$

4) Une autre méthode pour calculer $g(x)$ consiste à écrire la fonction $t\mapsto e^{-xt}\sin\sqrt{t}$ sous forme d'une série.

Pour cela, on va utiliser le développement en série entière de la fonction sin, qui donne, pour tout $t\geqslant 0$:

$$\sin\sqrt{t} = \sum_{n=0}^{+\infty}\frac{(-1)^n}{(2n+1)!}t^{n+1/2}.$$

✓ Il n'était pas possible ici d'utiliser le développement en série entière de e^{-xt} qui conduisait à des intégrales de $t^n\sin\sqrt{t}$ qui sont divergentes !

On a donc, pour tout $(x,t)\in\mathbb{R}_+^*\times\mathbb{R}_+$:

$$\begin{aligned}e^{-xt}\sin\sqrt{t} &= \sum_{n=0}^{+\infty}\frac{(-1)^n}{(2n+1)!}t^{n+1/2}e^{-xt} \\ &= \sum_{n=0}^{+\infty} u_n(t)\end{aligned}$$

en ayant posé, pour $x>0$ fixé, $u_n(t) = \dfrac{(-1)^n}{(2n+1)!}t^{n+1/2}e^{-xt}$.

Pour justifier l'interversion série-intégrale, nous allons ici utiliser le théorème d'intégration terme à terme.

On considère donc ici la série de fonctions $\sum\limits_{n\in\mathbb{N}} u_n$; les fonctions u_n sont évidemment continues sur $\mathbb{R}_+$, et le calcul fait ci-dessus montre que la série $\sum\limits_{n\in\mathbb{N}} u_n$ converge simplement sur $\mathbb{R}_+$ vers la fonction $t\mapsto e^{-xt}\sin\sqrt{t}$, qui est continue sur $\mathbb{R}_+$.

Les u_n sont bien intégrables sur $\mathbb{R}_+$ car

$$\begin{aligned}\lim_{t\to+\infty} t^2u_n(t) &= \lim_{t\to+\infty}\frac{(-1)^n}{(2n+1)!}t^{n+5/2}e^{-xt} \\ &= 0 \text{ (car } x>0)\,,\end{aligned}$$

donc $u_n(t) \underset{t\to+\infty}{=} o\left(\dfrac{1}{t^2}\right)$ (argument déjà utilisé). On a ensuite :

$$\begin{aligned}\int_{\mathbb{R}_+}|u_n| &= \frac{1}{(2n+1)!}\int_0^{+\infty} t^{n+1/2}e^{-xt}\,dt \\ &= \frac{1}{(2n+1)!}\cdot\frac{1}{x^{n+3/2}}\int_0^{+\infty} v^{n+1/2}e^{-v}\,dv,\end{aligned}$$

à l'aide du changement de variable $v=xt$. Nous allons calculer les intégrales :

$$I_n = \int_0^{+\infty} v^{n+1/2}e^{-v}\,dv$$

par récurrence sur n. Pour cela, on fait une intégration par parties, qui s'écrit :

$$\begin{aligned}I_{n+1} &= \int_0^{+\infty} v^{n+3/2}e^{-v}\,dv \\ &= \underbrace{\left[-v^{n+3/2}e^{-v}\right]_0^{+\infty}}_{=0} + \left(n+\frac{3}{2}\right)I_n\,.\end{aligned}$$

On a donc, pour $n\geqslant 1$, la relation :

$$I_n = \left(n+\frac{1}{2}\right)I_{n-1}\,,$$

puis, par itération de cette formule :

$$\begin{aligned} I_n &= \left(n+\frac{1}{2}\right) I_{n-1} \\ &= \left(n+\frac{1}{2}\right)\left(n-\frac{1}{2}\right) I_{n-2} \\ &= \cdots \\ &= \left(n+\frac{1}{2}\right)\left(n-\frac{1}{2}\right)\cdots\left(\frac{3}{2}\right) I_0 \\ &= \frac{(2n+1)(2n-1)\cdots 3}{2^n} I_0 \\ &= \frac{(2n+1)!}{2^n(2n)(2n-2)\cdots 2} I_0 \\ &= \frac{(2n+1)!}{2^{2n} n!} I_0 \,, \end{aligned}$$

ce qui donne finalement :

$$\begin{aligned} \int_{\mathbb{R}_+} |u_n| &= \frac{1}{x^{n+3/2}} \cdot \frac{1}{2^{2n} n!} I_0 \\ &= \frac{1}{x^{3/2}} \cdot \frac{1}{(4x)^n n!} I_0 \,. \end{aligned}$$

La série de terme général $\int_{\mathbb{R}_+} |u_n|$ est donc convergente, car on reconnaît ici, à une constante près, le développement en série entière de $\mathrm{e}^{\frac{1}{4x}}$.

Il résulte alors du théorème d'intégration terme à terme que :

$$\begin{aligned} \int_0^{+\infty} \mathrm{e}^{-xt} \sin\sqrt{t}\,\mathrm{d}t &= \sum_{n=0}^{+\infty} \int_{\mathbb{R}_+} u_n(t)\,\mathrm{d}t \\ &= \frac{1}{x^{3/2}} I_0 \sum_{n=0}^{+\infty} \frac{(-1)^n}{(4x)^n n!} = \frac{1}{x^{3/2}} I_0 \mathrm{e}^{\frac{-1}{4x}} \,. \end{aligned}$$

Enfin,

$$\begin{aligned} I_0 &= \int_0^{+\infty} \sqrt{v}\,\mathrm{e}^{-v}\,\mathrm{d}v \\ &= \frac{1}{2}\int_0^{+\infty} \frac{\mathrm{e}^{-v}}{\sqrt{v}}\,\mathrm{d}v \ \text{(intégration par parties)} \\ &= \int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t \quad (v=t^2) \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{1}{2} f(0) = \frac{\sqrt{\pi}}{2} \,, \end{aligned}$$

et l'on retrouve ainsi le résultat obtenu à la troisième question.

Exercice 14.9

1) On applique directement le théorème sur la continuité d'une intégrale à paramètre.

Posons, pour $(x,t) \in \mathbb{R} \times \mathbb{R}_+^*$,

$$f(x,t) = \frac{\sin xt}{\mathrm{e}^t - 1} \cdot$$

- Pour tout $x \in \mathbb{R}$, la fonction $t \mapsto f(x,t)$ est continue sur $\mathbb{R}_+^*$ par les théorèmes usuels.
- De même, pour tout $t > 0$, la fonction $x \mapsto f(x,t)$ est continue sur $\mathbb{R}$.
- Soit $a > 0$. Pour tout $x \in [-a\,;a]$, on a, pour tout $t > 0$ et compte tenu de l'inégalité $|\sin X| \leqslant |X|$:

$$|f(x,t)| \leqslant \frac{|xt|}{\mathrm{e}^t - 1} \leqslant a\frac{t}{\mathrm{e}^t - 1} = \varphi(t)\,,$$

et φ est continue sur $\mathbb{R}_+^*$ et intégrable, puisque $\lim\limits_{t\to 0} \varphi(t) = a$ et $\varphi(t) \underset{+\infty}{=} o\big(\frac{1}{t^2}\big)$.

Le théorème du cours permet de conclure que F est continue sur tout segment $[-a\,;a]$ donc sur $\mathbb{R}$.

2) • En utilisant le développement en série entière de $\dfrac{1}{1-X}$ pour $|X| < 1$, on a, pour tout $x \in \mathbb{R}$ et tout $t > 0$:

$$\begin{aligned} \frac{\sin xt}{\mathrm{e}^t - 1} &= \frac{\mathrm{e}^{-t}\sin xt}{1-\mathrm{e}^{-t}} = \mathrm{e}^{-t}\sin xt \sum_{n=0}^{+\infty} (\mathrm{e}^{-t})^n \\ &= \sum_{n=1}^{+\infty} u_n(t)\,, \end{aligned}$$

avec $u_n(t) = \mathrm{e}^{-nt}\sin xt$ pour $n \geqslant 1$.

• L'intégrale $I_n = \int_0^{+\infty} u_n$ est absolument convergente puisque $|u_n(t)| \leqslant \mathrm{e}^{-t}$ pour tout $t > 0$ et $n \geqslant 1$, et l'on a :

$$\begin{aligned} I_n &= \mathcal{I}m\left(\int_0^{+\infty} \mathrm{e}^{-t(n-\mathrm{i}x)}\,\mathrm{d}t\right) \\ &= \mathcal{I}m\left(\left[\frac{\mathrm{e}^{-t(n-\mathrm{i}x)}}{-n+\mathrm{i}x}\right]_0^{+\infty}\right) \\ &= \mathcal{I}m\left(\frac{1}{n-\mathrm{i}x}\right) = \frac{x}{x^2+n^2}\,, \end{aligned}$$

en ayant utilisé :

$$\lim_{t\to+\infty} \mathrm{e}^{-t(n-\mathrm{i}x)} = \lim_{t\to+\infty} \mathrm{e}^{-nt} \underbrace{\mathrm{e}^{\mathrm{i}xt}}_{\text{bornée}} = 0\,.$$

• Pour obtenir le résultat demandé, il s'agit donc d'intervertir $\sum$ et $\int$. On pourrait procéder comme dans l'exemple de **[SF14.23]** page 564, en utilisant le théorème de convergence dominée appliqué aux séries, mais on peut aussi faire une démonstration directe très simple.

Pour cela, posons, pour $n \in \mathbb{N}^*$:

$$r_n = \int_0^{+\infty} f(x,t)\,\mathrm{d}t - \sum_{k=1}^{n} I_k\,.$$

Alors :

$$\begin{aligned} r_n &= \int_0^{+\infty} \left(f(x,t) - \sum_{k=1}^{n} \mathrm{e}^{-kt} \sin xt \right) \mathrm{d}t \\ &= \int_0^{+\infty} \left(\sum_{k=n+1}^{+\infty} \mathrm{e}^{-kt} \sin xt \right) \mathrm{d}t \\ &= \int_0^{+\infty} \mathrm{e}^{-(n+1)t} \frac{\sin xt}{1-\mathrm{e}^{-t}}\,\mathrm{d}t \\ &= \int_0^{+\infty} \mathrm{e}^{-nt} f(x,t)\,\mathrm{d}t\,. \end{aligned}$$

Or pour $x \in \mathbb{R}$ fixé, la fonction $t \mapsto f(x,t)$ est bornée sur $\mathbb{R}_+$ (car elle est continue sur $\mathbb{R}_+^*$, et possède des limites finies en 0^+ et en $+\infty$). En notant :

$$M_x = \sup_{t \in \mathbb{R}_+} |f(x,t)|\ ,$$

on obtient : $|r_n| \leqslant M_x \displaystyle\int_0^{+\infty} \mathrm{e}^{-nt}\,\mathrm{d}t = \frac{M_x}{n}$.
On en déduit $\lim\limits_{n\to+\infty} r_n = 0$, ce qui prouve :

$$\int_0^{+\infty} f(x,t)\,\mathrm{d}t = \sum_{n=1}^{+\infty} I_n = \sum_{n=1}^{+\infty} \frac{x}{x^2+n^2}\,.$$

3) Soit $x > 0$ fixé.
La fonction $\varphi_x : u \mapsto \dfrac{x}{x^2+u^2}$ étant continue décroissante sur $\mathbb{R}_+$, on a, pour tout $n \in \mathbb{N}^*$:

$$\int_n^{n+1} \varphi_x(u)\,\mathrm{d}u \leqslant \varphi_x(n) \leqslant \int_{n-1}^{n} \varphi_x(u)\,\mathrm{d}u\,,$$

donc en sommant, la série et les intégrales considérées étant convergentes :

$$\int_1^{+\infty} \varphi_x \leqslant \sum_{n=1}^{+\infty} \frac{x}{x^2+n^2} \leqslant \int_0^{+\infty} \varphi_x\,.$$

Or $\int \dfrac{x}{x^2+u^2}\,\mathrm{d}u = \mathrm{Arctan}\left(\dfrac{u}{x}\right)$ donc finalement, pour tout $x > 0$:

$$\frac{\pi}{2} - \mathrm{Arctan}\,\frac{1}{x} \leqslant F(x) \leqslant \frac{\pi}{2}\,.$$

On en déduit : $\lim\limits_{x\to+\infty} F(x) = \dfrac{\pi}{2}$.

✓ On peut remarquer que :

$$\lim_{x\to+\infty} \sum_{n=1}^{+\infty} \frac{x}{x^2+n^2} \neq \sum_{n=1}^{+\infty} \lim_{x\to+\infty} \frac{x}{x^2+n^2}\,.$$

Le théorème de la double limite ne pouvait donc pas s'appliquer ici !

✓ Vous remarquerez encore une fois l'utilité de la méthode de comparaison série-intégrale.

Exercice 14.10

1) D'abord, on remarque que, pour tout $n \in \mathbb{N}^*$, la fonction $f_n : x \mapsto \dfrac{1}{\mathrm{ch}^n x}$ est continue sur $\mathbb{R}_+$. Le problème de l'intégrabilité de f_n ne se pose donc qu'au voisinage de $+\infty$.
Puisque $\mathrm{ch}\,x = \dfrac{\mathrm{e}^x + \mathrm{e}^{-x}}{2}$, on a $\mathrm{ch}\,x \underset{x\to+\infty}{\sim} \dfrac{\mathrm{e}^x}{2}$, donc $f_n(x) \underset{x\to+\infty}{\sim} 2^n \mathrm{e}^{-nx}$.
La fonction $x \mapsto \mathrm{e}^{-nx}$ étant positive et intégrable au voisinage de $+\infty$, il en est de même de f_n.

En conclusion, u_n est bien défini pour tout $n \in \mathbb{N}^*$.

2) On pense immédiatement ici à utiliser le théorème de convergence dominée, puisque $u_n = \displaystyle\int_{\mathbb{R}_+} f_n$.

– Les f_n sont continues et intégrables sur $\mathbb{R}_+$, comme on vient de le voir.

– On a, pour tout $n \in \mathbb{N}^*$, $f_n(0) = 1$, et, pour tout $x > 0$, $\lim\limits_{n\to+\infty} f_n(x) = 0$ puisque $0 < \dfrac{1}{\mathrm{ch}\,x} < 1$.
La suite (f_n) converge donc simplement sur $\mathbb{R}_+$ vers la fonction f définie par :

$$f(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x > 0 \end{cases}\ ,$$

et f est bien continue par morceaux sur $\mathbb{R}_+$.

– Il reste l'hypothèse de domination. Sa vérification est facile ici ; en effet, puisque $0 < \dfrac{1}{\mathrm{ch}\,x} \leqslant 1$, on a, pour tout $n \in \mathbb{N}^*$,

$$\left|\frac{1}{\mathrm{ch}^n x}\right| = \frac{1}{\mathrm{ch}^n x} \leqslant \frac{1}{\mathrm{ch}\,x}\,.$$

La fonction $x \mapsto \dfrac{1}{\mathrm{ch}\,x}$ étant intégrable sur $\mathbb{R}_+$ (cela correspond à l'étude de u_n dans le cas $n = 1$), l'hypothèse de domination est bien vérifiée.

Le théorème de convergence dominée s'applique et donne :

$$\lim_{n\to+\infty} u_n = \lim_{n\to+\infty} \int_{\mathbb{R}_+} f_n = \int_{\mathbb{R}_+} f = 0\,.$$

3) a) Pour étudier la série de terme général $u_n = \displaystyle\int_{\mathbb{R}_+} f_n$, commençons par étudier la série de terme général f_n.

Pour tout $x > 0$, on a $0 < \dfrac{1}{\operatorname{ch} x} < 1$, donc la série géométrique $\sum f_n(x)$ converge et a pour somme :

$$\sum_{n=1}^{+\infty} f_n(x) = \frac{1}{\operatorname{ch} x} \cdot \frac{1}{1-\frac{1}{\operatorname{ch} x}} = \frac{1}{\operatorname{ch} x - 1}\cdot$$

Or, la fonction $x \mapsto \dfrac{1}{\operatorname{ch} x - 1}$ n'est *pas* intégrable sur $\mathbb{R}_+^*$: en effet, $\operatorname{ch} x - 1 \underset{x\to 0}{\sim} \dfrac{x^2}{2}$, donc $\dfrac{1}{\operatorname{ch} x - 1} \underset{0}{\sim} \dfrac{2}{x^2}$, et la fonction $x \mapsto \dfrac{2}{x^2}$ n'est pas intégrable au voisinage de 0...

On peut cependant utiliser le théorème d'intégration terme à terme, bien qu'il ne s'applique pas directement ici, en raisonnant par l'absurde.

En effet, si la série de terme général $u_n = \displaystyle\int_{\mathbb{R}_+} f_n = \int_{\mathbb{R}_+} |f_n|$ était convergente, ce théorème affirme que la fonction $\displaystyle\sum_{n=1}^{+\infty} f_n$, qui est définie et continue sur $\mathbb{R}_+^*$, serait intégrable sur $\mathbb{R}_+^*$. Or on vient de voir que ce n'est pas le cas.
Cela montre donc, par l'absurde, la divergence de la série de terme général u_n.

b) • La convergence de la série de terme général $(-1)^n u_n$ résulte directement du critère spécial sur les séries alternées ; en effet :

- on a déjà vu que la suite (u_n) est positive, de limite nulle ;
- puisque, pour tout $x \geqslant 0$, $0 < \dfrac{\operatorname{ch} x}{\leqslant} 1$, on a, pour tout $n \in \mathbb{N}^*$, $\dfrac{1}{\operatorname{ch}^{n+1} x} \leqslant \dfrac{1}{\operatorname{ch}^n x}$, ce qui donne en intégrant : $u_{n+1} \leqslant u_n$.
  La suite (u_n) est donc décroissante.

Le critère spécial sur les séries alternées permet donc de conclure que la série de terme général $(-1)^n u_n$ converge.

• Pour calculer sa somme, nous allons faire une interversion série-intégrale. Or d'après ce qui a été fait auparavant, le théorème d'intégration terme à terme ne peut s'appliquer.

Nous allons donc utiliser le théorème de convergence dominée appliqué à la suite formée des sommes partielles de la série $\displaystyle\sum_{n\geqslant 1} (-1)^n f_n$.

On considère donc la série de fonctions $\displaystyle\sum_{n\geqslant 1} (-1)^n f_n$, et on notera, pour $n \in \mathbb{N}^*$, S_n la somme partielle : $S_n = \displaystyle\sum_{k=1}^{n} (-1)^k f_k$.

Vérifions les hypothèses du théorème.

- Les fonctions $(-1)^n f_n$ sont continues sur $[0\,;+\infty[$.
  Pour $x = 0$, $(-1)^n f_n(0) = (-1)^n$, donc la série $\displaystyle\sum_{n\geqslant 1} (-1)^n f_n(0)$ diverge.
  Pour $x > 0$, la série $\displaystyle\sum_{n\geqslant 1} (-1)^n f_n(x)$ est une série géométrique de raison $-\dfrac{1}{\operatorname{ch} x}$, donc elle converge puisque $-1 < -\dfrac{1}{\operatorname{ch} x} < 1 < 0$, et l'on a :
  $$\sum_{n=1}^{+\infty} \frac{(-1)^n}{\operatorname{ch}^n x} = -\frac{1}{\operatorname{ch} x} \cdot \frac{1}{1+\frac{1}{\operatorname{ch} x}} = -\frac{1}{\operatorname{ch} x + 1}\cdot$$
  La série de fonctions $\displaystyle\sum_{n\geqslant 1} (-1)^n f_n$ converge donc simplement sur $]0\,;+\infty[$ vers la fonction $S : x \mapsto \dfrac{-1}{\operatorname{ch} x + 1}$, qui est continue sur $\mathbb{R}_+^*$.
- *Hypothèse de domination :*
  Il faut ici majorer les S_n (en valeur absolue). Or, pour tout $x > 0$:
  $$S_n(x) = \sum_{k=1}^{n} \frac{(-1)^k}{\operatorname{ch}^k x}$$
  $$= -\frac{1}{\operatorname{ch} x} \cdot \frac{1-\frac{(-1)^n}{\operatorname{ch}^n x}}{1+\frac{1}{\operatorname{ch} x}} = -\frac{1-\frac{(-1)^n}{\operatorname{ch}^n x}}{\operatorname{ch} x + 1}\cdot$$
  Donc, pour tout $x > 0$ et tout $n \in \mathbb{N}^*$,
  $$|S_n(x)| = \frac{\left|1-\frac{(-1)^n}{\operatorname{ch}^n x}\right|}{\operatorname{ch} x + 1} \leqslant \frac{1+\frac{1}{\operatorname{ch}^n x}}{\operatorname{ch} x + 1} \leqslant \frac{2}{\operatorname{ch} x + 1}\cdot$$

La fonction $\varphi\colon x \mapsto \dfrac{2}{\operatorname{ch} x+1}$ est continue et intégrable sur $\mathbb{R}_+$ (pour des raisons tout à fait semblables à celles déjà vues), donc l'hypothèse de domination est vérifiée.

Le théorème de convergence dominée permet donc d'intervertir somme et intégrale, d'où :

$$\sum_{n=1}^{+\infty}(-1)^n u_n = \sum_{n=1}^{+\infty}\int_0^{+\infty}(-1)^n f_n$$
$$= \int_0^{+\infty} S = -\int_0^{+\infty}\frac{1}{\operatorname{ch} x+1}\,\mathrm{d}x\,.$$

Il ne reste plus qu'à calculer cette intégrale ce que nous faisons par les méthodes classiques exposées en **[SF10.6]**.

$$\int_0^{+\infty}\frac{1}{\operatorname{ch} x+1}\,\mathrm{d}x = \int_0^{+\infty}\frac{2\,\mathrm{d}x}{\mathrm{e}^x+\mathrm{e}^{-x}+2}$$
$$= \int_1^{+\infty}\frac{2\,\mathrm{d}t}{t\left(t+\frac{1}{t}+2\right)} \quad (t=\mathrm{e}^x)$$
$$= \int_1^{+\infty}\frac{2\,\mathrm{d}t}{(t+1)^2} = \left[\frac{-2}{t+1}\right]_1^{+\infty} = 1$$

et pour conclure :

$$\sum_{n=1}^{+\infty}(-1)^n u_n = -1\,.$$

Exercice 15.1

On considère les évènements :

$A\cap B$ de probabilité x ;

$A\cap\overline{B}$ de probabilité y ;

$\overline{A}\cap B$ de probabilité z ;

$\overline{A}\cap\overline{B}$ de probabilité t .

Puisque :

$$(A\cap B)\cup(A\cap\overline{B})\cup(\overline{A}\cap B)\cup(\overline{A}\cap\overline{B})$$
$$= \Big(A\cap(\underbrace{B\cup\overline{B}}_{=\Omega})\Big)\cup\Big(\overline{A}\cap(\underbrace{B\cup\overline{B}}_{=\Omega})\Big)$$
$$= A\cup\overline{A} = \Omega\,,$$

et que ces 4 évènements sont deux à deux incompatibles (car, par exemple,

$$(A\cap B)\cap(A\cap\overline{B}) = A\cap\underbrace{B\cap\overline{B}}_{=\emptyset} = \emptyset),$$

on a :

$$x+y+z+t=1\,.$$

D'autre part,

$$P(A) = P\big((A\cap B)\cup(A\cap\overline{B})\big) = x+y$$

et de même : $P(B)=x+z$.
On a donc :

$$P(A\cap B)-P(A)P(B)$$
$$= x-(x+y)(x+z)$$
$$= x(1-x-y-z)-yz$$
$$= xt-yz\,.$$

En utilisant alors la propriété :

$$\forall\, a\in[0\,;1],\ a(1-a)\leqslant\frac{1}{4}\,,$$

on a :

$$0\leqslant xt\leqslant x(1-x)\leqslant\frac{1}{4}\,,$$
$$0\leqslant yz\leqslant y(1-y)\leqslant\frac{1}{4}\,,$$

donc :

$$-\frac{1}{4}\leqslant P(A\cap B)-P(A)P(B)\leqslant\frac{1}{4}\,,$$

ce qui s'écrit aussi

$$|P(A\cap B)-P(A)P(B)|\leqslant\frac{1}{4}\cdot$$

Exercice 15.2

Supposons les billets numérotés de 1 à N. L'univers Ω, correspondant à tous les tirages possibles, est ici l'ensemble des parties de n éléments parmi N, donc $\operatorname{Card}(\Omega)=\binom{N}{n}$.
Si l'on désigne par G l'évènement :
« l'acheteur a au moins un billet gagnant »
alors l'évènement contraire est :
$\overline{G}$: « l'acheteur n'a eu aucun billet gagnant ».
$\overline{G}$ est donc formé des parties de n éléments parmi les $N-n$ tickets perdants, donc $\operatorname{Card}\overline{G}=\binom{N-n}{n}$ et la probabilité cherchée vaut :

$$P(G)=1-\frac{\binom{N-n}{n}}{\binom{N}{n}}\cdot$$

Exercice 15.3

1) D_d est l'ensemble des nombres de la forme dk avec $1\leqslant dk\leqslant n$ donc $1\leqslant k\leqslant\dfrac{n}{d}$.

Donc Card $D_d = \dfrac{n}{d}$, et puisque Card $\Omega = n$:

$$P(D_d) = \frac{1}{d}\cdot$$

2) a) Soit $s \in [\![2\,;r]\!]$. L'ensemble

$$D_{p_1} \cap D_{p_2} \cap \ldots \cap D_{p_s}$$

est l'ensemble des entiers de $[\![1\,;n]\!]$ qui sont multiples de $p_1, \ldots, p_s$, donc ce sont exactement ceux qui sont multiples du produit $p_1 p_2 \cdots p_s$ puisque les p_i sont premiers entre eux deux à deux.
Ainsi :

$$D_{p_1} \cap D_{p_2} \cap \ldots \cap D_{p_s} = D_{p_1 p_2 \cdots p_s}\,,$$

d'où

$$P\big(D_{p_1} \cap D_{p_2} \cap \ldots \cap D_{p_s}\big) = P\big(D_{p_1 p_2 \cdots p_s}\big)$$
$$= \frac{1}{p_1 p_2 \cdots p_s} = \prod_{i=1}^{s} \frac{1}{p_i} = \prod_{i=1}^{s} P(D_{p_i})\,,$$

ce qui prouve que les évènements D_{p_i} sont mutuellement indépendants.

b) Soit $A = \{k \in [\![1\,;n]\!] \mid k \wedge n = 1\}$.
Un entier k est premier avec n si et seulement si il n'admet aucun des p_i comme facteur premier. On a donc :

$$A = \overline{D_{p_1}} \cap \overline{D_{p_2}} \cap \ldots \cap \overline{D_{p_r}}.$$

Les évènements D_{p_i} étant indépendants, il en est de même de leurs complémentaires donc :

$$P(A) = \prod_{i=1}^{r} P\big(\overline{D_{p_i}}\big) = \prod_{i=1}^{r} \big(1 - P(D_{p_i})\big)$$
$$= \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

Or Card $A = \varphi(n)$ donc $P(A) = \dfrac{\varphi(n)}{n}$, et l'on retrouve ainsi la formule bien connue qui permet de calculer $\varphi(n)$.

Exercice 15.4

1) Notons C l'évènement : « un nombre fini seulement d'évènements A_k se réalise ». Cela signifie qu'à partir d'un certain rang n aucun des évènements ne se réalise, c'est-à-dire que l'on a $\bigcap_{k=n}^{+\infty} \overline{A_k}$. Ce rang n pouvant être quelconque, on a :

$$C = \bigcup_{n_0=0}^{+\infty} \left(\bigcap_{k=n_0}^{+\infty} \overline{A_k} \right) = \overline{B}$$

avec $B = \bigcap_{n=0}^{+\infty} \left(\bigcup_{k=n}^{+\infty} A_k \right)$.

Pour tout entier n, on a $B \subset \bigcup_{k=n}^{+\infty} A_n$, donc par sous-additivité :

$$\forall\, n \in \mathbb{N},\ P(B) \leqslant \sum_{k=n}^{+\infty} P(A_n)\,.$$

Mais $\sum_{k=n}^{+\infty} P(A_n)$ est le reste d'une série convergente donc tend vers 0 quand $n \to +\infty$ d'où $P(B) = 0$.

Par suite, $P(C) = 1$ et l'évènement C est quasi certain.

2) On a aussi, par sous-additivité de P :

$$P(C) \leqslant \sum_{n=0}^{+\infty} P\left(\bigcap_{k=n}^{+\infty} \overline{A_k} \right).$$

Calculons, pour tout n :

$$P\left(\bigcap_{k=n}^{+\infty} \overline{A_k} \right) = \lim_{N \to +\infty} P\left(\bigcap_{k=n}^{N} \overline{A_k} \right).$$

Les évènements A_k étant mutuellement indépendants, il en est de même des $\overline{A_k}$ donc :

$$P\left(\bigcap_{k=n}^{N} \overline{A_k} \right) = \prod_{k=n}^{N} P(\overline{A_k})$$
$$= \prod_{k=n}^{N} \big(1 - P(A_k)\big)$$
$$\leqslant \prod_{k=n}^{N} \mathrm{e}^{-P(A_k)} = \exp\left(-\sum_{k=n}^{N} P(A_k) \right)$$

à l'aide de l'inégalité $1 - x \leqslant \mathrm{e}^{-x}$.

Or par hypothèse $\lim_{N \to +\infty} \sum_{k=n}^{N} P(A_k) = +\infty$ (si une série à termes positifs diverge, c'est vers $+\infty$) donc :

$$P\left(\bigcap_{k=n}^{+\infty} \overline{A_k} \right) = \lim_{N \to +\infty} P\left(\bigcap_{k=n}^{N} \overline{A_k} \right) = 0\,.$$

On en déduit $P(C) = 0$: il est quasi impossible que seul un nombre fini d'évènements A_k se réalise, ce qui signifie qu'il est quasi certain qu'une infinité d'évènements se réalisent.

Exercice 15.5

Il faut évidemment supposer que la radio et le paysan font leurs prévisions indépendamment l'un de l'autre, et cela qu'il fasse beau ou pas.
Commençons par énumérer tous les évènements qui apparaissent dans l'énoncé :

B : « il fait beau » ;

M : « il fait mauvais » ;

R_M : « la radio annonce du mauvais temps » ;

P_B : « le paysan prédit du beau temps ».

Les hypothèses de l'énoncé s'écrivent alors :

$$P(B) = \frac{5}{7} \quad \text{et} \quad P(M) = \frac{2}{7}\cdot$$

$$P\left(R_M|B\right) = \frac{5}{100} \quad \text{et} \quad P\left(R_M|M\right) = \frac{95}{100}\cdot$$

$$P\left(P_B \mid B\right) = \frac{90}{100} \quad \text{et} \quad P\left(P_B \mid M\right) = \frac{10}{100}\cdot$$

Lorsque le paysan a prédit du beau temps et la radio de la pluie, cherchons la probabilité qu'il pleuve, c'est-à-dire

$$p = P\left(M \mid P_B \cap R_M\right).$$

Puisque $\{B,M\}$ est un système complet d'évènements, la formule de Bayes donne :

$$p = \frac{P(M \cap P_B \cap R_M)}{P(P_B \cap R_M)} = \frac{P(M \cap P_B \cap R_M)}{P(P_B \cap R_M \cap B)+P(P_B \cap R_M \cap M)}\cdot$$

Or :

$$P(P_B \cap R_M \cap B) = P\left(P_B \cap R_M|B\right) P(B)$$

et puisque les prédictions de la radio et du paysan sont indépendantes, on a :

$$P\left(P_B \cap R_M|B\right) = P\left(P_B|B\right) P\left(R_M|B\right).$$

Finalement :

$$\begin{aligned} P(P_B\cap R_M \cap B) &= P\left(P_B \mid B\right) P\left(R_M \mid B\right) P(B) \\ &= \frac{9}{10}\cdot\frac{5}{100}\cdot\frac{5}{7} \end{aligned}$$

et de la même façon :

$$\begin{aligned} P(P_B\cap R_M \cap M) &= P\left(P_B \mid M\right) P\left(R_M \mid M\right) P(M) \\ &= \frac{1}{10}\cdot\frac{95}{100}\cdot\frac{2}{7}\cdot \end{aligned}$$

Finalement :

$$p = \frac{\frac{1}{10}\cdot\frac{95}{100}\cdot\frac{2}{7}}{\frac{1}{10}\cdot\frac{95}{100}\cdot\frac{2}{7}+\frac{9}{10}\cdot\frac{5}{100}\cdot\frac{5}{7}} = \approx 0{,}46\,.$$

La probabilité qu'il pleuve est donc inférieure à 0,5 et Mr Lambda n'a pas besoin de prendre son parapluie. Finalement c'est le paysan qui avait raison, bien qu'il se trompe plus souvent. Sauriez-vous expliquer ce paradoxe ?

Exercice 15.6

1) Notons T_1 et T_2 les évènements :

T_k : « le tirage se fait dans l'urne n^{o} k » ($k \in \{1,2\}$).

$\{T_1,T_2\}$ est un système complet d'évènements équiprobables donc, par la formule des probabilités totales :

$$\begin{aligned} P(B_1) &= P\left(B_1 \mid T_1\right) P(T_1) \\ &\quad + P\left(B_2 \mid T_2\right) P(T_2) \\ &= \frac{1}{2}(p_1+p_2)\,. \end{aligned}$$

2) L'urne dans laquelle on pioche au n^{e} tirage est U_1 si l'évènement B_{n-1} est réalisé et U_2 si c'est $\overline{B_{n-1}}$. Donc, par la formule des probabilités totales :

$$\begin{aligned} P(B_n) &= P\left(B_n \mid B_{n-1}\right) P(B_{n-1}) \\ &\quad + P\left(B_n \mid \overline{B_{n-1}}\right) P(\overline{B_{n-1}}) \\ &= p_1 P(B_{n-1}) + p_2 P(\overline{B_{n-1}}) \\ &= p_1 P(B_{n-1}) + p_2(1-P(B_{n-1})) \\ &= (p_1-p_2)P(B_{n-1}) + p_2\,. \end{aligned}$$

3) La suite $\big(P(B_n)\big)$ est donc une suite arithmético-géométrique. On cherche le point fixe :

$$\ell = (p_1-p_2)\ell + p_2 \iff \ell = \frac{p_2}{p_2+1-p_1}$$

($p_2+1-p_1 \neq 0$ car $p_1,\, p_2 \in \left]0\,;1\right[$).
D'après le cours, la suite de terme général :

$$u_n = P(B_n) - \ell$$

est géométrique de raison p_1-p_2 donc

$$\forall\, n \in \mathbb{N}^*,\ u_n = (p_1-p_2)^{n-1} u_1$$

d'où :

$$P(B_n) = (p_1-p_2)^{n-1}\big[P(B_1)-\ell\big] + \ell\,.$$

Puisque p_1 et p_2 sont dans $]0\,;1[$, $|p_1 - p_2| < 1$ donc :

$$\lim_{n\to+\infty} P(B_n) = \ell = \frac{p_2}{p_2 + 1 - p_1}\cdot$$

Exercice 15.7

1) Pour tout $n \in \mathbb{N}$ on considère les évènements :

E_n : « la famille comporte n enfants ».

La famille $\{E_n, n \in \mathbb{N}\}$ est un système complet d'évènements incompatibles donc on doit avoir $\sum_{n=0}^{+\infty} p_n = 1$, ce qui donne :

$$p_0 = 1 - \sum_{n=1}^{+\infty} \alpha p^n = 1 - \alpha\frac{p}{1-p}$$

soit : $p_0 = \dfrac{1-p(1+\alpha)}{1-p}$.

✓ La condition de l'énoncé se traduit donc par $p_0 > 0$.

2) Soit $k \geqslant 1$ et les évènements :

G_k : « la famille comporte exactement k garçons »

D'après la formule des probabilités totales on a :

$$P(G_k) = \sum_{n=0}^{+\infty} P(G_k \cap E_n)\,.$$

Mais $G_k \cap E_n = \emptyset$ pour $k > n$ (pour avoir k garçons, il faut au moins k enfants !). Donc il reste :

$$\begin{aligned} P(G_k) &= \sum_{n=k}^{+\infty} P(G_k \cap E_n) \\ &= \sum_{n=k}^{+\infty} P(G_k \mid E_n)\, P(E_n)\,. \end{aligned}$$

Dans une famille de n enfants, la répartition des garçons suit la loi binomiale de paramètres n et $\frac{1}{2}$ donc :

$$P(G_k \mid E_n) = \binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k},$$

$$\begin{aligned} \text{d'où : } P(G_k) &= \sum_{n=k}^{+\infty}\binom{n}{k}\left(\frac{1}{2}\right)^n \alpha p^n \\ &= \frac{\alpha}{k!}\left(\frac{p}{2}\right)^k \sum_{n=k}^{+\infty}\frac{n!}{(n-k)!}\left(\frac{p}{2}\right)^{n-k}. \end{aligned}$$

Or en dérivant k fois le développement en série entière :

$$\forall x \in]-1\,;1[,\ \frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n\,,$$

on obtient, pour $|x| < 1$:

$$\frac{k!}{(1-x)^{k+1}} = \sum_{n=k}^{+\infty}\frac{n!}{(n-k)!}x^{n-k},$$

$$\begin{aligned} \text{d'où : } P(G_k) &= \alpha\left(\frac{p}{2}\right)^k \frac{1}{(1-\frac{p}{2})^{k+1}} \\ &= \frac{2\alpha p^k}{(2-p)^{k+1}}\,. \end{aligned}$$

3) L'évènement :

« G'_1 : la famille a au moins un garçon »

s'écrit :

$$G'_1 = \bigcup_{k=1}^{+\infty} G_k\,,$$

et l'évènement :

« G'_2 : la famille a au moins 2 garçons »

s'écrit :

$$G'_2 = \bigcup_{k=2}^{+\infty} G_k\,.$$

On nous demande de calculer :

$$P\left(G'_2 \mid G'_1\right) = \frac{P(G'_2 \cap G'_1)}{P(G'_1)} = \frac{P(G'_2)}{P(G'_1)}$$

puisque $G'_1 \cap G'_2 = G'_2$.

Les évènements de la famille $(G_k)_{k\geqslant 2}$ étant incompatibles :

$$\begin{aligned} P(G'_2) &= \sum_{k=2}^{+\infty} P(G_k) = \sum_{k=2}^{+\infty}\frac{2\alpha p^k}{(2-p)^{k+1}} \\ &= \frac{2\alpha p^2}{(2-p)^3}\sum_{k=2}^{+\infty}\left(\frac{p}{2-p}\right)^{k-2} \\ &= \frac{2\alpha p^2}{(2-p)^3}\cdot\frac{1}{1-\frac{p}{2-p}} \\ &= \frac{\alpha p^2}{(1-p)(2-p)^2}\,. \end{aligned}$$

Puis :

$$\begin{aligned} P(G'_1) &= P(G_1 \cup G'_2) = P(G_1) + P(G'_2) \\ &= \frac{2\alpha p}{(2-p)^2} + \frac{\alpha p^2}{(1-p)(2-p)^2} \\ &= \frac{\alpha p}{(1-p)(2-p)}\,. \end{aligned}$$

Finalement la probabilité demandée vaut :

$$P\left(G'_2 \mid G'_1\right) = \frac{P(G'_2)}{P(G'_1)} = \frac{p}{2-p}\,.$$

Exercice 15.8

On considèrera les évènements :

B_i : « le i^{e} tirage a donné une boule blanche » ;

R_i : « le i^{e} tirage a donné une boule rouge » ;

A : « les n premiers tirages ont donné exactement une boule blanche ».

1) L'évènement A est la réunion disjointe des évènements :

C_k : « le k^{e} tirage a donné une boule blanche et les autres tirages des rouges ».

Ainsi :

$$C_k = R_1 \cap \ldots \cap R_{k-1} \cap B_k \cap R_{k+1} \ldots \cap R_n$$

et par la formule des probabilités composées :

$$\begin{aligned} P(C_k) = P(R_1)P\left(R_2|R_1\right) \times \cdots \\ \times P\left(R_{k-1} \mid R_{k-2} \cap \ldots \cap R_1\right) \\ \times P\left(B_k \mid R_{k-1} \cap \ldots \cap R_1\right) \\ \times \cdots \times P\left(R_n \mid R_{n-1} \cap \ldots \cap R_1\right) \end{aligned}$$

donc :

$$\begin{aligned} P(C_k) &= \left(\frac{r}{b+r}\right)^{k-1} \frac{b}{b+r} \left(\frac{r}{b+r-1}\right)^{n-k} \\ &= \frac{br^{n-1}}{(b+r)^k(b+r-1)^{n-k}} \,. \end{aligned}$$

Les évènements C_k étant incompatibles, on en déduit :

$$\begin{aligned} P(A) &= \sum_{k=1}^{n} P(C_k) \\ &= \frac{br^{n-1}}{(b+r-1)^n} \sum_{k=1}^{n} \left(\frac{b+r-1}{b+r}\right)^k \\ &= \frac{br^{n-1}}{(b+r)(b+r-1)^{n-1}} \frac{1-\left(\frac{b+r-1}{b+r}\right)^n}{1-\frac{b+r-1}{b+r}} \\ &= \frac{br^{n-1}}{(b+r-1)^{n-1}} \left(1-\left(\frac{b+r-1}{b+r}\right)^n\right). \end{aligned}$$

2) Avec les notations précédentes, on cherche ici la probabilité $P\left(C_n \mid A\right)$.

Or :

$$P\left(C_n \mid A\right) = \frac{P(C_n \cap A)}{P(A)} = \frac{P(C_n)}{P(A)}$$

puisque $C_n \subset A$. Il ne reste plus qu'à remplacer par les quantités calculées auparavant pour trouver, après simplification :

$$P\left(C_n \mid A\right) = \frac{(b+r-1)^{n-1}}{(b+r)^n - (b+r-1)^n} \,.$$

Exercice 16.1

1) • La variable X est à valeurs dans $\mathbb{N}^*$ car le premier saut est toujours réussi.
Pour $n \in \mathbb{N}^*$, l'évènement $(X = n)$ signifie que le sauteur a réussi les n premiers sauts et a échoué au $(n+1)^{\text{e}}$. Les sauts étant indépendants on a :

$$\begin{aligned} P(X=n) &= 1 \cdot \frac{1}{2} \cdots \frac{1}{n} \cdot \left(1 - \frac{1}{n+1}\right) \\ &= \frac{1}{n!} - \frac{1}{(n+1)!} \,. \end{aligned}$$

Clairement, la série $\sum_{n \geqslant 1} P(X=n)$ est une série télescopique de somme 1.

2) Puisque $P(X=n) \leqslant \frac{1}{n!}$, la série de terme général $nP(X=n)$ converge (absolument) par le théorème de comparaison de séries à termes positifs c'est-à-dire que $E(X)$ existe. De plus :

$$\begin{aligned} E(X) &= \sum_{n=1}^{+\infty} n \left(\frac{1}{n!} - \frac{1}{(n+1)!}\right) \\ &= \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} - \sum_{n=1}^{+\infty} \frac{n}{(n+1)!} \\ &= \mathrm{e} - \sum_{n=1}^{+\infty} \frac{(n+1)-1}{(n+1)!} \\ &= \mathrm{e} - \sum_{n=1}^{+\infty} \left(\frac{1}{n!} - \frac{1}{(n+1)!}\right) \\ &= \mathrm{e} - 1 \quad \text{(télescopage)} \,. \end{aligned}$$

• Pour les mêmes raisons que ci-dessus, $E(X^2)$ existe et :

$$\begin{aligned} E(X^2) &= \sum_{n=1}^{+\infty} n^2 \frac{n}{(n+1)!} \\ &= \sum_{n=1}^{+\infty} (n^2 - 1 + 1) \frac{n}{(n+1)!} \\ &= \sum_{n=1}^{+\infty} \frac{n(n^2-1)}{(n+1)!} + \underbrace{\sum_{n=1}^{+\infty} \frac{n}{(n+1)!}}_{=1} \\ &= \sum_{n=2}^{+\infty} \frac{1}{(n-2)!} + 1 = \mathrm{e} + 1 \,. \end{aligned}$$

Par la formule de Kœnig-Huygens :

$$V(X) = E(X^2) - (E(X))^2$$
$$= \mathrm{e} + 1 - (\mathrm{e}-1)^2 = \mathrm{e}(3-\mathrm{e})\,.$$

Exercice 16.2

1) Pour tout $k \in \mathbb{N}^*$, notons Y_k la variable aléatoire égale au nombre de tablettes qu'il faut acheter, lorsqu'on a déjà $k-1$ images, pour en obtenir une différente. C'est en quelque sorte le temps d'attente entre deux images distinctes.

On a donc : $X_n = \sum_{k=1}^{n} Y_k$.

D'autre part, si $k-1$ images ont été obtenues, chaque nouvelle tablette a une probabilité $\dfrac{n-(k-1)}{n}$ de donner une nouvelle image. La variable Y_k correspond donc au rang du 1er « succès » dans une série d'épreuves indépendantes où un succès a pour probabilité $\dfrac{n-(k-1)}{n}$.

Par suite, Y_k suit la loi géométrique de paramètre $\dfrac{n-(k-1)}{n}$ d'où :

$$E(Y_k) = \frac{n}{n-(k-1)}\,.$$

Par linéarité de l'espérance, on a alors :

$$E(X_n) = \sum_{k=1}^{n} \frac{n}{n-(k-1)} = n\sum_{k=1}^{n} \frac{1}{k}\,.$$

Par un résultat connu sur les sommes partielles de la série harmonique, on en déduit :

$$E(X_n) \underset{n\to+\infty}{\sim} n\ln n\,.$$

2) • D'après le cours, la variance de chaque variable Y_k est égale à :

$$V(Y_k) = \frac{1-\frac{n-(k-1)}{n}}{\left(\frac{n-(k-1)}{n}\right)^2} = n\frac{k-1}{(n-(k-1))^2}\,.$$

Les variables Y_k étant indépendantes, on en déduit :

$$V(X_n) = \sum_{k=1}^{n} V(Y_k) = n\sum_{k=1}^{n} \frac{k-1}{(n-(k-1))^2}$$
$$\leqslant n^2 \sum_{k=1}^{n} \frac{1}{(n-(k-1))^2}$$
$$\leqslant n^2 \sum_{k=1}^{n} \frac{1}{k^2} \leqslant n^2 \sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6} n^2\,.$$

On a bien prouvé l'existence d'une constante $C > 0$ telle que :

$$V(X_n) \leqslant Cn^2\,.$$

• L'inégalité de Bienaymé-Tchebychev donne alors, pour tout $\varepsilon > 0$:

$$P\big(|X_n - E(X_n)| \geqslant \varepsilon n\ln n\big) \leqslant \frac{V(X_n)}{\varepsilon^2 n^2 (\ln n)^2} \leqslant \frac{C}{\varepsilon^2 (\ln n)^2}$$

ce qui entraîne $(*)$:

$$\lim_{n\to+\infty} P\left(|X_n - E(X_n)| \geqslant \varepsilon n\ln n\right) = 0\,.$$

D'autre part, on a vu que $E(X_n) \underset{n\to+\infty}{\sim} n\ln n$, ce qui implique :

$$\exists n_0 \in \mathbb{N},\ n \geqslant n_0 \implies \left|\frac{E(X_n)}{n\ln n} - 1\right| \leqslant \frac{\varepsilon}{2}\,.$$

Pour $n \geqslant n_0$ on aura donc :

$$\left|\frac{X_n}{n\ln n} - 1\right| \leqslant \left|\frac{X_n}{n\ln n} - \frac{E(X_n)}{n\ln n}\right| + \left|\frac{E(X_n)}{n\ln n} - 1\right|$$
$$\leqslant \left|\frac{X_n - E(X_n)}{n\ln n}\right| + \frac{\varepsilon}{2}$$

donc, pour $n \geqslant n_0$, si $\left|\dfrac{X_n}{n\ln n} - 1\right| \geqslant \varepsilon$ alors on aura :

$$\left|\frac{X_n - E(X_n)}{n\ln n}\right| \geqslant \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}\,,$$

d'où :

$$P\left(\left|\frac{X_n}{n\ln n} - 1\right| \geqslant \varepsilon\right) \leqslant P\left(|X_n - E(X_n)| \geqslant \frac{\varepsilon}{2} n\ln n\right)$$

puisque l'évènement de gauche implique celui de droite, donc est inclus dans celui-ci.

Compte tenu de $(*)$ on aura donc :

$$\lim_{n\to+\infty} P\left(\left|\frac{X_n}{n\ln n} - 1\right| \geqslant \varepsilon\right) = 0\,,$$

ce qui est le résultat voulu.

Exercice 16.3

Soit X (respectivement Y) la variable aléatoire égale au nombre d'œufs pondus par la poule (respectivement par la cane).

X et Y prennent des valeurs entières et, pour tout $(k,\ell) \in \mathbb{N}^2$ on a :

$$P(X=k) = \mathrm{e}^{-\alpha}\frac{\alpha^k}{k!}\ ,\ P(Y=\ell) = \mathrm{e}^{-\beta}\frac{\beta^\ell}{\ell!}\cdot$$

X et Y étant indépendantes, on sait que leur somme $Z = X+Y$ suit la loi de Poisson de paramètre $\alpha+\beta$.

La question posée est de calculer :

$$P\left(X=k \mid Z=n\right).$$

Or :

$$P\left(X=k \mid Z=n\right) = \frac{P\left(X=k \cap Z=n\right)}{P(Z=n)},$$

et :

$$\begin{aligned} P\left(X=k \cap Z=n\right) &= P\left(X=k \cap Y=n-k\right) \\ &= P(X=k)P(Y=n-k) \\ &= \mathrm{e}^{-\alpha}\frac{\alpha^k}{k!}\mathrm{e}^{-\beta}\frac{\beta^{n-k}}{(n-k)!}, \end{aligned}$$

donc finalement :

$$\begin{aligned} &P\left(X=k \mid Z=n\right) \\ &= \mathrm{e}^{-(\alpha+\beta)}\frac{\alpha^k}{k!}\frac{\beta^{n-k}}{(n-k)!}\cdot\frac{1}{\mathrm{e}^{-(\alpha+\beta)}\frac{(\alpha+\beta)^n}{n!}} \\ &= \binom{n}{k}\left(\frac{\alpha}{\alpha+\beta}\right)^k\left(\frac{\beta}{\alpha+\beta}\right)^{n-k}\cdot \end{aligned}$$

La loi conditionnelle de X sachant $(Z=n)$ est donc la loi binomiale de paramètres $\left(n, \frac{\alpha}{\alpha+\beta}\right)$.

Exercice 16.4

1) On a $X(\Omega) = [\![1\,;n]\!]$.
Soit $k \in [\![1\,;n]\!]$. Comme il y a en tout $1+2+\cdots+n = \frac{n(n+1)}{2}$ boules dans l'urne, dont exactement k numérotées « k », on a :

$$P(X=k) = \frac{k}{n(n+1)/2} = \frac{2k}{n(n+1)}\cdot$$

X est une variable aléatoire finie donc elle admet une espérance donnée par :

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} kP(X=k) \\ &= \sum_{k=1}^{n}\frac{2k^2}{n(n+1)} \\ &= \frac{2}{n(n+1)}\cdot\frac{n(n+1)(2n+1)}{6} \\ &= \frac{2n+1}{3}\cdot \end{aligned}$$

✓ Il faut connaître les formules donnant $\sum_{k=1}^{N} k$, $\sum_{k=1}^{N} k^2$ et $\sum_{k=1}^{N} k^3$, que nous avons déjà eu l'occasion de rappeler.

2) a) On a $X(\Omega) = Y(\Omega) = [\![1\,;n]\!]$. Soit $(i,j) \in [\![1\,;n]\!]^2$.

• Si $i \neq j$, on a :

$$\begin{aligned} &P\left(X=i \cap Y=j\right) \\ &\quad = P\left(Y=j \mid X=i\right)P(X=i) \\ &\quad = \frac{2i}{n(n+1)} \times \frac{j}{\frac{n(n+1)}{2}-1} \\ &\quad = \frac{4ij}{(n-1)n(n+1)(n+2)}\cdot \end{aligned}$$

• Si $i = j$, on a :

$$\begin{aligned} &P\left(X=i \cap Y=j\right) \\ &\quad = P\left(Y=j \mid X=i\right)P(X=i) \\ &\quad = \frac{2i}{n(n+1)} \times \frac{i-1}{\frac{n(n+1)}{2}-1} \\ &\quad = \frac{4i(i-1)}{(n-1)n(n+1)(n+2)}\cdot \end{aligned}$$

b) Soit $j \in [\![1\,;n]\!]$. D'après la formule des probabilités totales avec le système complet d'évènements $\{(X=i),\ i \in [\![1\,;n]\!]\}$ on a :

$$\begin{aligned} P(Y=j) &= \sum_{k=1}^{n} P\left(X=i \cap Y=j\right) \\ &= \sum_{\substack{i=1 \\ i\neq j}}^{n}\frac{4ij}{(n-1)n(n+1)(n+2)} \\ &\qquad + \frac{4j(j-1)}{(n-1)n(n+1)(n+2)} \\ &= \frac{4j}{(n-1)n(n+1)(n+2)} \\ &\qquad \times \left(\frac{n(n+1)}{2} - j\right) \\ &\qquad + \frac{4j(j-1)}{(n-1)n(n+1)(n+2)} \\ &= \frac{4j}{(n-1)n(n+1)(n+2)} \\ &\qquad \times \left(\frac{n(n+1)}{2} - j + j - 1\right) \\ &= \frac{2j}{n(n+1)}\cdot \end{aligned}$$

On constate que la loi de Y est la même que celle de X.

c) On a : $P\left(X=1\cap Y=1\right)=0$ et :

$$P(X=1)\times P(Y=1)=\frac{4}{n^2(n+1)^2}\,,$$

donc les variables X et Y ne sont pas indépendantes.

Exercice 16.5

1) Calculons d'abord, pour tout $i\in\mathbb{N}$:

$$\begin{aligned}
&\sum_{j=0}^{+\infty}P\left(X=i\cap Y=j\right)\\
&\quad=\sum_{j=0}^{+\infty}\frac{(i+j)\lambda^{i+j}}{\mathrm{e}\,i!j!}\\
&\quad=\frac{\lambda^i}{\mathrm{e}\,i!}\left(i\sum_{j=0}^{+\infty}\frac{\lambda^j}{j!}+\sum_{j=0}^{+\infty}j\frac{\lambda^j}{j!}\right)\\
&\quad=\frac{\lambda^i}{\mathrm{e}\,i!}\left(i\sum_{j=0}^{+\infty}\frac{\lambda^j}{j!}+\sum_{j=1}^{+\infty}\lambda\frac{\lambda^{j-1}}{(j-1)!}\right)\\
&\quad=\frac{\lambda^i}{\mathrm{e}\,i!}(i+\lambda)\mathrm{e}^{\lambda}\,.
\end{aligned}$$

Puis :

$$\begin{aligned}
&\sum_{i=0}^{+\infty}\frac{\lambda^i}{\mathrm{e}\,i!}(i+\lambda)\mathrm{e}^{\lambda}\\
&\quad=\frac{\mathrm{e}^{\lambda}}{\mathrm{e}}\left(\sum_{i=1}^{+\infty}\lambda\frac{\lambda^{i-1}}{(i-1)!}+\lambda\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}\right)\\
&\quad=\frac{\mathrm{e}^{\lambda}}{\mathrm{e}}\left(2\lambda\mathrm{e}^{\lambda}\right)=2\lambda\mathrm{e}^{2\lambda-1}\,.
\end{aligned}$$

Cela prouve déjà la sommabilité de la série double $\sum\limits_{(i,j)\in\mathbb{N}^2}P\left(X=i\cap Y=j\right)$, mais il faut de plus que la somme de cette série soit égale à 1 c'est-à-dire :

$$2\lambda\mathrm{e}^{2\lambda-1}=1\,.$$

La fonction $\lambda\mapsto 2\lambda\mathrm{e}^{2\lambda-1}-1$ sur étant strictement croissante sur $\mathbb{R}_+$ ne peut s'y annuler qu'une fois, et c'est pour $\lambda=\dfrac{1}{2}$.

2) • En réutilisant le calcul fait ci-dessus, on obtient, pour tout $i\in\mathbb{N}$:

$$\begin{aligned}
P(X=i)&=\sum_{j=0}^{+\infty}P\left(X=i\cap Y=j\right)\\
&=\frac{\lambda^i}{\mathrm{e}\,i!}(i+\lambda)\mathrm{e}^{\lambda}\\
&=\frac{1}{2^i i!}\mathrm{e}^{-1/2}\left(i+\frac{1}{2}\right).
\end{aligned}$$

• On en déduit immédiatement, par symétrie, que pour tout $j\in\mathbb{N}$ on a :

$$P(Y=j)=\frac{1}{2^j j!}\mathrm{e}^{-1/2}\left(j+\frac{1}{2}\right).$$

3) On a $P\left(X=0\cap Y=0\right)=0$ alors que

$$P(X=0)P(Y=0)\neq 0\,,$$

donc les variables X et Y ne sont pas indépendantes.

4) Par le théorème de transfert, et sous réserve d'absolue convergence, on a

$$\begin{aligned}
&E\left(2^{X+Y}\right)\\
&\quad=\sum_{(i,j)\in\mathbb{N}^2}2^{i+j}P\left(X=i\cap Y=j\right).
\end{aligned}$$

On calcule d'abord, pour tout $i\in\mathbb{N}$:

$$\begin{aligned}
&\sum_{j=0}^{+\infty}2^{i+j}P\left(X=i\cap Y=j\right)\\
&\quad=\sum_{j=0}^{+\infty}2^{i+j}\frac{i+j}{\mathrm{e}\,i!j!}\left(\frac{1}{2}\right)^{i+j}\\
&\quad=\frac{1}{\mathrm{e}\,i!}\left(\sum_{j=0}^{+\infty}\frac{i}{j!}+\sum_{j=0}^{+\infty}\frac{j}{j!}\right)\\
&\quad=\frac{1}{\mathrm{e}\,i!}(i+1)\mathrm{e}=\frac{i+1}{i!}\,.
\end{aligned}$$

Ce calcul montre en même temps la convergence, donc la convergence absolue, de la série.

Puisque la série $\sum\limits_i\dfrac{i+1}{i!}$ converge (absolument), on peut affirmer à l'aide du théorème de Fubini **[S11.14]** que $E\left(2^{X+Y}\right)$ existe et que :

$$E\left(2^{X+Y}\right)=\sum_{i=0}^{+\infty}\frac{i+1}{i!}=2\mathrm{e}\,.$$

Exercice 16.6

1) • On a $X_i(\Omega)=[\![0\,;\min(a,n)]\!]\subset[\![0\,;n]\!]$.

Calculons $P(X_i=k)$ pour $0\leqslant k\leqslant n$.

Le nombre de tirages possibles dans l'urne est $\binom{aN}{n}$.

Un tirage réalise l'évènement $(X_i=k)$ s'il est formé de k boules de couleur i exactement et donc de $n-k$ d'une autre couleur. Le nombre de façons de choisir les

k boules de couleur i est égal à $\binom{a}{k}$ (qui est nul par convention si $k > a$), et le nombre de façons de choisir $n-k$ boules d'une autre couleur est égal à $\binom{aN-a}{n-k}$.

On a donc :

$$P(X_i = k) = \frac{\binom{a}{k}\binom{a(N-1)}{n-k}}{\binom{aN}{n}}.$$

Une telle loi s'appelle une loi hypergéométrique.

• *Calcul de l'espérance*

X_i ne prenant qu'un nombre fini de valeurs, son espérance existe et vaut :

$$E(X_i) = \sum_{k=0}^{n} k \frac{\binom{a}{k}\binom{a(N-1)}{n-k}}{\binom{aN}{n}}.$$

$$= \binom{aN}{n}^{-1} \sum_{k=1}^{n} \underbrace{k\binom{a}{k}}_{} \binom{a(N-1)}{n-k}$$

$$= \binom{aN}{n}^{-1} \sum_{k=1}^{n} \overbrace{a\binom{a-1}{k-1}} \binom{a(N-1)}{n-k}$$

$$= \binom{aN}{n}^{-1} a \sum_{k=0}^{n-1} \binom{a-1}{k}\binom{a(N-1)}{(n-1)-k}$$

$$= \binom{aN}{n}^{-1} a\binom{aN-1}{n-1}$$

(à l'aide de la formule de Vandermonde)

$$= \left[\frac{aN}{n}\binom{aN-1}{n-1}\right]^{-1} a\binom{aN-1}{n-1}.$$

Finalement : $E(X_i) = \dfrac{n}{N}$.

• *Calcul de la variance*

Nous calculons d'abord $E(X_i(X_i-1))$, de la même façon que ci-dessus.

$$E(X_i(X_i-1)) = \sum_{k=0}^{n} k(k-1)P(X_i = k)$$

$$= \binom{aN}{n}^{-1} \sum_{k=2}^{n} \underbrace{k(k-1)\binom{a}{k}} \binom{a(N-1)}{n-k}$$

$$= \binom{aN}{n}^{-1} \sum_{k=2}^{n} \overbrace{a(a-1)\binom{a-2}{k-2}} \binom{a(N-1)}{n-k}$$

$$= a(a-1)\binom{aN}{n}^{-1} \sum_{k=0}^{n-2} \binom{a-2}{k}\binom{a(N-1)}{n-2-k}$$

$$= a(a-1)\binom{aN}{n}^{-1}\binom{aN}{n}\binom{aN-2}{n-2}$$

(à l'aide de la formule de Vandermonde)

$$= \frac{a(a-1)}{\frac{aN}{n}\frac{aN-1}{n-1}\binom{aN-2}{n-2}}\binom{aN-2}{n-2}.$$

Finalement :

$$E(X_i(X_i-1) = \frac{(a-1)n(n-1)}{N(aN-1)}.$$

On termine le calcul grâce à la formule de Kœnig-Huyghens :

$V(X_i) = E(X_i(X_i-1)) + E(X_i) - \big[E(X_i)\big]^2$.

$$V(X_i) = \frac{(a-1)n(n-1)}{N(aN-1)} + \frac{n}{N} - \frac{n^2}{N^2}$$

$$= \frac{n}{N}\left(1 - \frac{1}{N}\right)\frac{aN-n}{aN-1}.$$

✓ Retenez qu'il est souvent plus facile de calculer $E(X(X-1))$ que $E(X^2)$ directement.

✓ Ces calculs sont assez fastidieux, et exigent de bien connaître les propriétés des coefficients binomiaux rappelés en **[S15.1]**.

2) Puisque $X_1 + \cdots + X_N$ est le nombre total de boules tirées, on a bien :

$$X_1 + X_2 + \cdots + X_N = n.$$

Pour tout $i \in [\![1\,;n]\!]$ on a :

$$\operatorname{cov}(X_1 + \cdots + X_N, X_i) = V(X_i) + \sum_{j \neq i} \operatorname{cov}(X_i, X_j)$$

et puisque la variable aléatoire $X_1 + \cdots + X_N$ est constante, $\operatorname{cov}(X_1 + \cdots + X_N, X_i) = 0$.

D'autre part, puisque toutes les couleurs jouent des rôles symétriques, les covariances $\operatorname{cov}(X_i,X_j)$ pour $i \neq j$ sont toutes égales. Finalement on obtient, pour $i \neq j$:

$$(N-1)\operatorname{cov}(X_i,X_j) + V(X_i) = 0$$

ce qui permet de trouver :

$$\operatorname{cov}(X_i,X_j) = -\frac{n}{N^2}\,\frac{aN-n}{aN-1}\,,$$

puis :

$$\rho(X_i,X_j) = \frac{\operatorname{cov}(X_i,X_j)}{\sigma(X_i)\sigma(X_j)} = \frac{\operatorname{cov}(X_i,X_j)}{V(X_i)} = -\frac{1}{N-1}\,.$$

✓ Le fait de trouver un coefficient de corrélation négatif signifie que les variables aléatoires X_i et X_j ont tendance à varier en sens contraires. Cela est normal puisque la somme des X_i est constante égale à n.

Exercice 16.7

Les fonctions génératrices de X_1 et X_2 sont les fonctions polynomiales :

$$G_{X_1}(t) = p_1 t + \cdots + p_6 t^6\,;$$
$$G_{X_2}(t) = q_1 t + \cdots + q_6 t^6\,.$$

Si le problème proposé avait une solution, $X_1 + X_2$ suivrait la loi uniforme sur $[\![2\,;12]\!]$ donc on aurait :

$$G_{X_1+X_2}(t) = \frac{1}{11}(t^2 + \cdots + t^{12}) = \frac{t^2}{11}\,\frac{1-t^{11}}{1-t}$$

pour tout réel $t \neq 1$.
Les lancers des deux dés étant indépendants, on sait que :

$$G_{X_1+X_2} = G_{X_1} G_{X_2}\,.$$

On devrait donc avoir :

$$(p_1 + p_2 t + \cdots + p_6 t^5) \times (q_1 + q_2 t + \cdots + q_6 t^5) = \frac{1}{11}\,\frac{1-t^{11}}{1-t}\,.$$

Mais chaque polynôme du membre gauche de cette égalité étant de degré impair admet au moins une racine réelle alors que le polynôme du second membre n'admet que des racines complexes non réelles, d'où la contradiction.

Exercice 16.8

Désignons par C_i pour $1 \leqslant i \leqslant N$ l'évènement :

« le i^{e} champignon cueilli est comestible ».

Par la formule des probabilités totales appliquée avec le système complet d'évènements $\{(N=n),\, n \in \mathbb{N}^*\}$, la probabilité $P(C)$ que tous les champignons cueillis soient comestibles est :

$$P(C) = \sum_{n=1}^{+\infty} P(C_1 \cap \ldots \cap C_n \mid N=n) \times P(N=n)\,.$$

Il faut évidemment supposer les évènements C_i indépendants entre eux et indépendants de N, de sorte que :

$$P(C_1 \cap \ldots \cap C_n \mid N=n) = P(C_1 \cap \ldots \cap C_n) = \prod_{i=1}^{n} P(C_i) = p^n\,.$$

On obtient donc :

$$P(C) = \sum_{n=1}^{+\infty} P(N=n)p^n = G_N(p)\,.$$

Exercice 16.9

Cet exercice généralise le précédent.

1) Par la formule des probabilités totales on a, pour tout $n \in \mathbb{N}$:

$$P(S=n) = \sum_{k=0}^{+\infty} P\big(X_1 + \cdots + X_k = n \cap N = k\big) = \sum_{k=0}^{+\infty} P(X_1 + \cdots + X_k = n) P(N=k)\,.$$

En effet, les variables aléatoires N, X_1, $X_2, \ldots$ étant supposés indépendantes, il en est de même de $X_1 + \cdots + X_k$ et de N.
Soit $t \in [-1\,;1]$. Puisque la série $\sum_{n\in\mathbb{N}} P(S=n)t^n$ est absolument convergente, le théorème de Fubini permet

d'écrire :

$$\begin{aligned} &G_S(t) \\ &= \sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} P\left(\sum_{i=1}^{k} X_i = n\right) P(N=k)t^n \\ &= \sum_{k=0}^{+\infty} P(N=k) \sum_{n=0}^{+\infty} P\left(\sum_{i=1}^{k} X_i = n\right) t^n \\ &= \sum_{k=0}^{+\infty} P(N=k) G_{X_1+\cdots+X_k}(t)\,. \end{aligned}$$

Les variables aléatoires X_i étant indépendantes et de même loi, on a :

$$G_{X_1+\ldots+X_k}(t) = \prod_{i=1}^{k} G_{X_i}(t) = G_X(t)^k\,,$$

ce qui donne :

$$\forall\, t \in [-1\,;1],\ G_S(t) = G_N\big[G_X(t)\big]\,.$$

2) • Lorsque les rayons de convergence des fonctions génératrices sont strictement supérieurs à 1, on sait que G_N et G_X sont dérivables en 1 donc G_S aussi (dérivabilité de la composée) et l'on a :

$$\begin{aligned} G'_S(1) &= G'_X(1)G'_N\big[G_X(1)\big] \\ &= G'_X(1)G'_N(1)\,, \end{aligned}$$

c'est-à-dire :

$$E(S) = E(X)E(N)\,.$$

• C'est un peu plus délicat si l'on ne sait rien sur les rayons de convergence. Dans ce cas, le fait que les espérances de N et X existent prouve seulement que les dérivées *à gauche* des fonctions génératrices existent. Mais il n'y a pas dans le cas général de résultat sur la dérivée à gauche d'une fonction composée !

On remarque cependant que, puisque $G'_X(1^-) = E(X) > 0$ (car X est à valeurs entières positives), G_X est strictement croissante au voisinage de 1^- donc, puisque $G_X(1) = 1$, on a $G_X(t) < 1$ dans un voisinage à gauche de 1. On peut alors écrire, pour t dans ce voisinage :

$$(*)\quad \begin{aligned} \frac{G_S(t)-G_S(1)}{t-1} &= \frac{G_N\big[G_X(t)\big]-1}{t-1} \\ &= \frac{G_N\big[G_X(t)\big]-1}{G_X(t)-1} \times \frac{G_X(t)-1}{t-1}\,, \end{aligned}$$

et puisque

$$\lim_{t\to 1^-} G_X(t) = G_X(1)^- = 1^-\,,$$

(en effet, la série qui définit une fonction génératrice étant normalement convergente sur $[-1\,;1]$, une fonction génératrice est forcément continue sur cet intervalle), on obtient en faisant tendre t vers 1^- dans $(*)$, que G_S est dérivable à gauche en 1 et que sa dérivée à gauche est encore donnée par :

$$G'_S(1^-) = G'_X(1^-)G'_N(1^-)\,.$$

On obtient donc le même résultat qu'auparavant.

Exercice 17.1

• On résout l'équation séparément sur $\mathbb{R}_-^*$ et $\mathbb{R}_+^*$ puisque le coefficient de y' s'annule en 0.
Sur chacun de ces intervalles l'équation homogène s'écrit :

$$y' = -\frac{\mathrm{e}^x}{\mathrm{e}^x-1}\cdot$$

Puisque $\displaystyle\int \frac{\mathrm{e}^x}{\mathrm{e}^x-1}\,\mathrm{d}x = \ln|\mathrm{e}^x-1|$, la solution générale de l'équation homogène est :

$$y = \frac{C}{\mathrm{e}^x-1}\cdot$$

Pour résoudre l'équation avec second membre, on utilise la méthode de variation de la constante en posant :

$$y(x) = \frac{C(x)}{\mathrm{e}^x-1}$$

qui conduit à :

$$C'(x) = 1, \quad \text{soit} \quad C(x) = x+\lambda\,.$$

La solution générale de l'équation est donc :

$$y(x) = \begin{cases} \dfrac{x+\lambda}{\mathrm{e}^x-1} & \text{si } x>0 \\[2ex] \dfrac{x+\mu}{\mathrm{e}^x-1} & \text{si } x<0 \end{cases}$$

avec λ, μ constantes réelles.

Puisque $\mathrm{e}^x - 1 \underset{x\to 0}{\sim} x$, la seule solution possible sur $\mathbb{R}$ est :

$$f(x) = \frac{x}{\mathrm{e}^x-1} \text{ si } x\neq 0\,,\ f(0)=1\,.$$

Cette solution convient bien : en effet la fonction $g\colon x \mapsto \dfrac{\mathrm{e}^x-1}{x}$, prolongée en 0 par $g(0)=1$, est développable en série entière de rayon de convergence $+\infty$ (il suffit d'écrire le développement de e^x) donc est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$. Puisque g ne s'annule pas, la fonction $f = \dfrac{1}{g}$ est aussi

de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$, donc c'est bien une solution sur $\mathbb{R}$ de l'équation proposée.

Exercice 17.2

• Nous allons d'abord résoudre l'équation.
L'équation homogène est : $y' + ay = 0$; c'est une équation à coefficients constants, admettant pour solutions les fonctions de la forme $x \mapsto Ce^{-ax}$, où C est une constante. La fonction $x \mapsto e^{-ax}$ étant de classe $\mathscr{C}^\infty$ et ne s'annulant pas, on peut chercher les solutions de (E) sous la forme : $y(x) = z(x)e^{-ax}$.
L'équation équivaut alors à :

$$z'(x)e^{-ax} = f(x) \quad \text{soit } z'(x) = e^{ax} f(x)\,,$$

puis $z(x) = \displaystyle\int_0^x e^{at} f(t)\,dt + C$.
On en déduit que l'ensemble des solutions de (E) est l'ensemble des fonctions définies sur $\mathbb{R}$ de la forme :

$$y(x) = \left(\int_0^x e^{at} f(t)\,dt + C\right) e^{-ax}$$

où C est une constante complexe quelconque.

> ✓ Ce résultat figure directement dans le cours, mais comme il est souvent oublié, nous avons préféré le redémontrer.

• Si une solution y est 2π-périodique, alors nécessairement $y(0) = y(2\pi)$. Pour éviter des calculs pénibles, démontrons la réciproque.
Soit y une solution de (E). Alors la fonction $z : x \mapsto y(x + 2\pi)$ est encore solution de (E) puisque f est 2π-périodique. Si y vérifie $y(0) = y(2\pi)$, alors $y(0) = z(0)$: y et z sont deux solutions de (E) qui coïncident en un point, elles sont donc égales d'après le théorème de Cauchy-Lipschitz. On a donc $y(x) = z(x)$ pour tout x, soit $y(x) = y(x + 2\pi)$, et y est 2π-périodique.
On a donc démontré le résultat préliminaire suivant :
y solution de (E) est 2π-périodique si et seulement si $y(0) = y(2\pi)$.
On peut maintenant résoudre les questions proprement dites.

1) Soit y une solution de E écrite sous la forme :

$$y(x) = \left(\int_0^x e^{at} f(t)\,dt + C\right) e^{-ax}\,.$$

L'égalité $y(0) = y(2\pi)$ équivaut à :

$$C = e^{-2a\pi}\left(\int_0^{2\pi} e^{at} f(t)\,dt + C\right)$$

soit à :

$$(1 - e^{-2a\pi})C = e^{-2a\pi}\int_0^{2\pi} e^{at} f(t)\,dt \quad (1)$$

Or, si $a \notin i\mathbb{Z}$, $e^{-2a\pi}$ est différent de 1.

En effet, en posant $z = x + iy$, $(x,y) \in \mathbb{R}^2$, l'équation $e^z = 1$ s'écrit $e^x(\cos y + i \sin y) = 1$, donc équivaut au système :

$$\begin{cases} e^x \cos y = 1 \\ e^x \sin y = 0 \end{cases}$$

soit à $x = 0$ et $y \in 2\pi\mathbb{Z}$.
Ainsi : $e^z = 1 \iff z \in 2i\pi\mathbb{Z}$.

Ici, $-2a\pi$ n'appartient pas à $2i\pi\mathbb{Z}$ puisque $a \notin i\mathbb{Z}$, donc $(1 - e^{-2a\pi}) \neq 0$ et l'équation (1) possède une et une seule solution C (peu importe sa valeur).
Conclusion : si $a \notin i\mathbb{Z}$, l'équation (E) admet une et une seule solution 2π-périodique.

2) • Déjà, le calcul fait ci-dessus (relation (1)) montre que, s'il existe une solution 2π-périodique y lorsque $a \in i\mathbb{Z}$, puisque $1 - e^{-2a\pi} = 0$, on doit nécessairement avoir $\displaystyle\int_0^{2\pi} e^{at} f(t)\,dt = 0$.
• Supposons réciproquement cette condition remplie.
On reprend la formule qui donne la forme générale des solutions :

$$y(x) = \left(\int_0^x e^{at} f(t)\,dt + C\right) e^{-ax}\,.$$

Puisqu'ici, $a \in i\mathbb{Z}$, la fonction $x \mapsto e^{-ax}$ est 2π-périodique. Le produit et le quotient de fonctions 2π-périodiques étant aussi 2π-périodique, il en résulte que y est 2π-périodique si et seulement si la fonction :

$$x \mapsto \int_0^x e^{at} f(t)\,dt + C$$

l'est, c'est-à-dire si la fonction $g : x \mapsto \displaystyle\int_0^x e^{at} f(t)\,dt$ l'est.
Or : $g(x + 2\pi) - g(x)$

$$= \int_0^{x+2\pi} e^{at} f(t)\,dt - \int_0^x e^{at} f(t)\,dt$$

$$= \int_x^{x+2\pi} e^{at} f(t)\,dt\,,$$

et, puisque la fonction $t \mapsto \mathrm{e}^{at} f(t)$ est 2π-périodique (comme produit de telles fonctions), on en déduit que :

$$g(x+2\pi) - g(x) = \int_0^{2\pi} \mathrm{e}^{at} f(t)\,\mathrm{d}t = 0\,,$$

c'est-à-dire que g est 2π-périodique.

On a donc bien démontré l'équivalence souhaitée.

Exercice 17.3

$\mathcal{M}_{n,1}(\mathbb{R})$ est ici muni du produit scalaire canonique défini par :

$$\forall\, X,\, Y \in \mathcal{M}_{n,1}(\mathbb{R}),\ \langle X \,|\, Y\rangle = {}^{\mathrm{t}}XY\,.$$

• Soit X une solution du système différentiel $X' = AX$.
On a, en utilisant les formules de dérivation :

$$\big(B(f,g)\big)' = B(f',g) + B(f,g')$$

lorsque B est une application bilinéaire (on l'applique ici au produit scalaire), et aussi la formule :

$$(u \circ f)' = u \circ f'$$

lorsque u est une application linéaire (ici la transposition) :

$$({}^{\mathrm{t}}XX)' = {}^{\mathrm{t}}X'X + {}^{\mathrm{t}}XX' = {}^{\mathrm{t}}X({}^{\mathrm{t}}A + A)X\,.$$

• Ainsi, si A est antisymétrique, $({}^{\mathrm{t}}XX)' = 0$ et X est de norme constante.

• Réciproquement, supposons que chaque solution du système différentiel $X' = AX$ est de norme constante. Alors, en reprenant le calcul précédent, pour toute telle solution X, on a :

$$\forall\, t \in \mathbb{R},\ X(t)({}^{\mathrm{t}}A + A)X(t) = 0\,.$$

Mais d'après le théorème de Cauchy-Lipschitz, t_0 étant donné dans $\mathbb{R}$ et X_0 donné dans $\mathcal{M}_{n,1}(\mathbb{R})$, il existe une solution du système différentiel telle que $X(t_0) = X_0$.

On en déduit que :

$$\forall\, X_0 \in \mathcal{M}_{n,1}(\mathbb{R}),\ {}^{\mathrm{t}}X_0({}^{\mathrm{t}}A + A)X_0 = 0\,.$$

Si λ est une valeur propre de ${}^{\mathrm{t}}A + A$ et si X_0 en est un vecteur propre associé, on aura alors :

$$\lambda\, {}^{\mathrm{t}}X_0 X_0 = 0 \text{ soit } \lambda \|X_0\|^2 = 0$$

d'où $\lambda = 0$.
0 est la donc seule valeur propre de la matrice symétrique réelle ${}^{\mathrm{t}}A + A$; celles-ci étant diagonalisable, on obtient ${}^{\mathrm{t}}A + A = 0$, c'est-à-dire A antisymétrique.

Exercice 17.4

• On introduit ici un changement de variable de la forme $x = \varphi(t)$, où φ doit être une application de classe $\mathscr{C}^2$ d'un certain intervalle I sur $\mathbb{R}_+^*$ dont la dérivée ne s'annule pas. En posant $z(t) = y\big(\varphi(t)\big)$ pour $t > 0$, on a :

$$\begin{aligned} z'(t) &= \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\big(y(x)\big) \\ &= \frac{\mathrm{d}y}{\mathrm{d}x}.\frac{\mathrm{d}x}{\mathrm{d}t} = y'(x).\varphi'(t)\,, \end{aligned}$$

puis

$$\begin{aligned} z''(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\big(y'(x).\varphi'(t)\big) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big(y'(x)\big)\varphi'(t) + y'(x).\varphi''(t) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\big(y'(x)\big)\frac{\mathrm{d}x}{\mathrm{d}t}\varphi'(t) + y'(x).\varphi''(t) \\ &= y''(x)\varphi'^2(t) + y'(x)\varphi''(t)\,, \end{aligned}$$

soit encore :

$$y'(x) = \frac{z'(t)}{\varphi'(t)}$$

et

$$y''(x) = \frac{z''(t)}{\varphi'^2(t)} - \frac{z'(t)\varphi''(t)}{\varphi'^3(t)}\,,$$

tout cela ayant bien un sens car φ' ne s'annule pas.

✓ Nous avons ici utilisé les notations différentielles pour faire les calculs, à la manière des physiciens. Une autre écriture, que certains trouveront peut-être plus rigoureuse, est la suivante :
$z'(t) = (y \circ \varphi)'(t) = \varphi'(t) \cdot (y' \circ \varphi)(t)$
puis

$$\begin{aligned} z''(t) &= \big(\varphi' \cdot (y' \circ \varphi)\big)'(t) \\ &= \varphi''(t) \cdot (y' \circ \varphi)(t) \\ &\qquad + \big(\varphi'(t)\big)^2 \cdot (y'' \circ \varphi)(t)\,. \end{aligned}$$

En remplaçant alors x par $\varphi(t)$ et $y(x)$, $y'(x)$ et $y''(x)$ par ces valeurs dans l'équation différentielle homogène :

$$(E_0) \quad 2xy'' + y' - y = 0\,,$$

on obtient l'équation différentielle équivalente :

$$2\varphi(t)\left(\frac{z''(t)}{\varphi'^2(t)} - \frac{z'(t)\varphi''(t)}{\varphi'^3(t)}\right) + \frac{z'(t)}{\varphi'(t)} - z(t) = 0$$

soit

$$2\frac{\varphi(t)}{\varphi'^2(t)}z''(t) + \frac{\varphi'^2(t) - 2\varphi(t)\varphi''(t)}{\varphi'^3(t)}z'(t) - z(t) = 0\,.$$

Pour que le coefficient de $z''(t)$ soit constant, il suffit que l'on ait, en supposant φ et φ' à valeurs strictement positives sur I, $\frac{\varphi'(t)}{\sqrt{\varphi(t)}} = cste$ d'où φ de la forme $t \mapsto (at+b)^2$, avec $a > 0$.
On a alors :

$$\varphi'^2(t) - 2\varphi(t)\varphi''(t) = 4a^2(at+b)^2 - 2a^2.2(at+b)^2 = 0\,,$$

donc le coefficient de $z'(t)$ est également nul. On peut alors choisir a et b de sorte que l'équation obtenue soit la plus simple possible, par exemple $2\frac{\varphi(t)}{\varphi'^2(t)} = 1$, ce qui donne $\varphi(t) = \frac{t^2}{2}$.
Réciproquement, ce changement de variable est bien une bijection de classe $\mathscr{C}^2$ de $\mathbb{R}_+^*$ sur $\mathbb{R}_+^*$, et il conduit à l'équation équivalente :

$$(E') \quad z''(t) - z(t) = t^2 + \frac{1}{\operatorname{sh} t}\,.$$

• La résolution de l'équation homogène associée ne pose pas de problème.
Pour chercher une solution particulière, nous allons utiliser le principe de superposition de solutions.

– On cherche d'abord une solution particulière à l'équation :

$$(E_1') \quad z''(t) - z(t) = t^2\,.$$

D'après le cours, il existe une telle solution particulière qui est un polynôme en t de degré 2 ; en cherchant z sous la forme $z(t) = at^2 + bt + c$, on trouve facilement $z(t) = -t^2 - 2$.

– Puis on cherche une solution particulière de l'équation

$$(E_2') \quad z''(t) - z(t) = \frac{1}{\operatorname{sh} t}\,.$$

Pour cela, on utilise la méthode de variations des constantes.
Puisque $y_1 : t \mapsto \operatorname{ch} t$ et $y_2 : t \mapsto \operatorname{sh} t$ sont deux solutions linéairement indépendantes de l'équation homogène associée, on cherche z sous la forme $z = \lambda_1 y_1 + \lambda_2 y_2$, où λ_1 et λ_2 sont deux fonctions de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ telles que :

$$\begin{cases} \lambda_1' y_1 + \lambda_2' y_2 = 0 \\ \lambda_1' y_1' + \lambda_2' y_2' = \dfrac{1}{\operatorname{sh} t} \end{cases}$$

soit

$$\begin{cases} \lambda_1' \operatorname{ch} t + \lambda_2' \operatorname{sh} t = 0 \\ \lambda_1' \operatorname{sh} t + \lambda_2' \operatorname{ch} t = \dfrac{1}{\operatorname{sh} t} \end{cases}.$$

Le déterminant de ce système est $\operatorname{ch}^2 t - \operatorname{sh}^2 t = 1$ donc les formules de Cramer donnent directement :

$$\lambda_1' = \begin{vmatrix} 0 & \operatorname{sh} t \\ \frac{1}{\operatorname{sh} t} & \operatorname{ch} t \end{vmatrix} = -1$$

$$\text{et} \quad \lambda_2' = \begin{vmatrix} \operatorname{ch} t & 0 \\ \operatorname{sh} t & \frac{1}{\operatorname{sh} t} \end{vmatrix} = \frac{\operatorname{ch} t}{\operatorname{sh} t}$$

d'où $\lambda_1(t) = -t$ et $\lambda_2(t) = \ln(\operatorname{sh} t)$ (nous n'avons pas écrit les constantes d'intégration puisque nous ne cherchons qu'une solution particulière).
Nous obtenons donc comme solution de (E_2') : $z(t) = -t\operatorname{ch} t + \operatorname{sh} t \ln(\operatorname{sh} t)$.

– Le principe de superposition de solutions donne alors comme solution particulière de (E') :

$$z(t) = -t^2 - 2 - t\operatorname{ch} t + \operatorname{sh} t \ln(\operatorname{sh} t)\,.$$

Pour obtenir l'ensemble des solutions de (E'), il faut alors ajouter à cette solution particulière l'ensemble des solutions de l'équation homogène associée, c'est-à-dire des fonctions de la forme $t \mapsto \lambda \operatorname{ch} t + \mu \operatorname{sh} t$.
On obtient alors finalement :

l'ensemble des solutions sur $\mathbb{R}_+^*$ de l'équation différentielle :

$$2xy'' + y' - y = 2x + \frac{1}{\operatorname{sh}(\sqrt{2x})}$$

est l'ensemble des applications de la forme :

$$x \mapsto \lambda \operatorname{ch}(\sqrt{2x}) + \mu \operatorname{sh}(\sqrt{2x}) - 2x - 2 - \sqrt{2x}\operatorname{ch}(\sqrt{2x}) + \operatorname{sh}(\sqrt{2x})\ln(\operatorname{sh}(\sqrt{2x}))$$

avec $(\lambda,\mu) \in \mathbb{R}^2$.

Exercice 17.5

On vérifie sans peine que la fonction $x \mapsto \mathrm{e}^x$ est bien solution de l'équation homogène $(x+1)y'' - y' - xy = 0$.

Cette fonction ne s'annulant pas sur $\mathbb{R}$, on peut donc chercher les solutions de (E) sous la forme $y(x) = \mathrm{e}^x z(x)$, où la fonction z sera deux fois dérivable sur le même domaine que y (c'est la méthode de variation de la constante).

On a alors :

$$y = \mathrm{e}^x z \quad , \quad y' = \mathrm{e}^x(z + z')$$
$$\text{et} \quad y'' = \mathrm{e}^x(z + 2z' + z'')$$

donc (E) est équivalente à :

$$(x+1)\mathrm{e}^x(z + 2z' + z'') - \mathrm{e}^x(z + z') - x\mathrm{e}^x z = \mathrm{e}^{-x}$$

ou encore à :

$$(x+1)z'' + (2x+1)z' = \mathrm{e}^{-2x} \quad (E')$$

En posant $u = z'$, (E') est donc une équation différentielle linéaire du premier ordre :

$$(x+1)u' + (2x+1)u = \mathrm{e}^{-2x} \quad (E')$$

que l'on résout par la méthode habituelle.

• *Supposons dans un premier temps que* $x \in]-1\,;+\infty[$, *soit* $x+1 > 0$

L'équation homogène associée à (E') est :

$$(x+1)u' + (2x+1)u = 0 \quad (E'_0)$$

dont on peut directement écrire les solutions :

$$u(x) = C\exp\left(-\int \frac{2x+1}{x+1}\,\mathrm{d}x\right)$$

avec C constante.

En écrivant $2x+1 = 2(x+1)-1$, on obtient :

$$\int \frac{2x+1}{x+1}\,\mathrm{d}x = \int \left(2 - \frac{1}{x+1}\right)\mathrm{d}x$$
$$= 2x - \ln|x+1|$$
$$= 2x - \ln(x+1)\,,$$

d'où la solution générale de (E'_0) :

$$u(x) = C\exp(-2x + \ln(x+1))$$
$$= C(x+1)\mathrm{e}^{-2x}\,.$$

Pour résoudre l'équation avec second membre (E'), on utilise la méthode de variation de la constante, en cherchant u sous la forme $u(x) = C(x)(x+1)\mathrm{e}^{-2x}$, où C est une fonction dérivable sur $]-1\,;+\infty[$ (cela est possible puisque $(x+1)\mathrm{e}^{-2x}$ ne s'annule pas sur cet intervalle).

En remplaçant dans (E'), on obtient :

$$C'(x)(x+1)^2\mathrm{e}^{-2x} = \mathrm{e}^{-2x}$$

On aboutit donc à : $C'(x) = \dfrac{1}{(x+1)^2}$ d'où $C(x) = -\dfrac{1}{x+1} + \lambda$, où λ est une constante.

Il ne reste plus qu'à remplacer dans tous les calculs précédents ; on obtient :

$$z'(x) = u(x) = \left(-\frac{1}{x+1} + \lambda\right)(x+1)\mathrm{e}^{-2x}$$
$$= -\mathrm{e}^{-2x} + \lambda(x+1)\mathrm{e}^{-2x}\,,$$

d'où :

$$z(x) = \frac{1}{2}\mathrm{e}^{-2x} + \lambda\int (x+1)\mathrm{e}^{-2x}\,\mathrm{d}x\,.$$

On calcule la primitive précédente à l'aide d'une intégration par parties :

$$\int \underbrace{(x+1)}_{u(x)}\underbrace{\mathrm{e}^{-2x}}_{v'(x)}\mathrm{d}x$$
$$= -\frac{(x+1)\mathrm{e}^{-2x}}{2} + \frac{1}{2}\int \mathrm{e}^{-2x}\,\mathrm{d}x$$
$$= -\frac{(x+1)\mathrm{e}^{-2x}}{2} - \frac{\mathrm{e}^{-2x}}{4} + cste$$
$$= -\frac{2x+3}{4}\mathrm{e}^{-2x} + cste\,,$$

ce qui conduit, quitte à changer la constante λ, à :

$$z(x) = \frac{1}{2}\mathrm{e}^{-2x} + \lambda(2x+3)\mathrm{e}^{-2x} + \mu$$

avec λ, μ constantes.

On obtient alors la solution générale de l'équation initiale, sur $]-1\,;+\infty[$, en multipliant cette expression par e^x, ce qui donne :

$$y(x) = \frac{1}{2}\mathrm{e}^{-x} + \lambda(2x+3)\mathrm{e}^{-x} + \mu\mathrm{e}^x$$

avec λ, μ constantes.

• *Cas où* $x \in]-\infty\,;-1[$, *soit* $x+1 < 0$

Les calculs sont similaires ; la seule différence est dans la primitivation de $\dfrac{1}{x+1}$ qui donne $\ln|x+1|$, de sorte que l'on obtient

$$u(x) = C\exp(-2x + \ln|x+1|)$$
$$= C|x+1|\,\mathrm{e}^{-2x}\,.$$

Mais, quitte à changer C en $-C$, on peut conserver la même expression qu'auparavant, et le reste des calculs est identique.

Bien sûr, les constantes d'intégration obtenues ne sont pas forcément les mêmes que dans le cas précédent, de sorte que l'ensemble des solutions de l'équation initiale sur $\mathbb{R}\setminus\{-1\}$ est l'ensemble des fonctions de la forme :

$$y(x)=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-x}+\lambda_1(2x+3)\mathrm{e}^{-x}+\mu_1\mathrm{e}^{x}\\ \qquad\text{pour } x\in\,]-\infty\,;-1[\\ \dfrac{1}{2}\mathrm{e}^{-x}+\lambda_2(2x+3)\mathrm{e}^{-x}+\mu_2\mathrm{e}^{x}\\ \qquad\text{pour } x\in\,]-1\,;+\infty[\end{cases}$$

où λ_1, λ_2, μ_1, μ_2 sont des constantes réelles.

• *Raccordement des solutions*

Pour terminer l'exercice, il faut chercher parmi les solutions précédentes celles qui sont solutions sur $\mathbb{R}$, ce qui exige de trouver parmi ces solutions celles qui sont prolongeables en une fonction deux fois dérivable sur $\mathbb{R}$.
Pour cela, il faut et il suffit que, si l'on note :

$$y_1(x)=\frac{1}{2}\mathrm{e}^{-x}+\lambda_1(2x+3)\mathrm{e}^{-x}+\mu_1\mathrm{e}^{x}\,,$$
$$y_2(x)=\frac{1}{2}\mathrm{e}^{-x}+\lambda_2(2x+3)\mathrm{e}^{-x}+\mu_2\mathrm{e}^{x}\,,$$

on ait les 3 relations

$$\lim_{x\to-1^-}y_1(x)=\lim_{x\to-1^+}y_2(x)\,,$$
$$\lim_{x\to-1^-}y_1'(x)=\lim_{x\to-1^+}y_2'(x)\,,$$
$$\lim_{x\to-1^-}y_1''(x)=\lim_{x\to-1^+}y_2''(x)\,.$$

Or, pour $i\in\{1,2\}$,

$$y_i'(x)=-\frac{1}{2}\mathrm{e}^{-x}+\lambda_i(-2x-1)\mathrm{e}^{-x}+\mu_i\mathrm{e}^{x}\,,$$
$$y_i''(x)=\frac{1}{2}\mathrm{e}^{-x}+\lambda_i(2x-1)\mathrm{e}^{-x}+\mu_i\mathrm{e}^{x}\,,$$

de sorte que les 3 conditions précédentes s'écrivent :

$$\begin{cases}\lambda_1\mathrm{e}+\dfrac{\mu_1}{\mathrm{e}} &= \lambda_2\mathrm{e}+\dfrac{\mu_2}{\mathrm{e}}\\ -3\lambda_1\mathrm{e}+\dfrac{\mu_1}{\mathrm{e}} &= -3\lambda_2\mathrm{e}+\dfrac{\mu_2}{\mathrm{e}}\,,\end{cases}$$

ce qui donne $\lambda_1=\lambda_2$ et $\mu_1=\mu_2$.

Réciproquement, toute fonction de la forme :

$$y(x)=\frac{1}{2}\mathrm{e}^{-x}+\lambda(2x+3)\mathrm{e}^{-x}+\mu\mathrm{e}^{x}$$

avec λ, μ constantes sera bien solution de l'équation sur $\mathbb{R}$: en effet, elle est deux fois dérivable sur $\mathbb{R}$, elle vérifie bien l'équation pour $x\in\,]-\infty\,;-1[$ et $x\in\,]-1\,;+\infty[$, et elle vérifie aussi l'équation pour $x=-1$ puisque l'égalité $(x+1)y''(x)-y'(x)-xy(x)=\mathrm{e}^{-x}$ étant vraie pour $x\neq-1$ reste vraie en $x=-1$ par passage à la limite et continuité.

En conclusion, les solutions de (E) sur $\mathbb{R}$ sont les fonctions de la forme

$$y(x)=\frac{1}{2}\mathrm{e}^{-x}+\lambda(2x+3)\mathrm{e}^{-x}+\mu\mathrm{e}^{x}$$

avec λ, μ constantes.

Exercice 17.6

Supposons qu'il existe une solution développable en série entière, sous la forme :

$$y(x)=\sum_{n=0}^{+\infty}a_nx^n\,,$$

et que cette série entière est de rayon de convergence $R>0$.
On a alors, pour $x\in\,]-R\,;R[$:

$$\begin{aligned}&x(1-x)y''(x)+(1-3x)y'(x)-y(x)\\ &=(x-x^2)\sum_{n=2}^{+\infty}n(n-1)a_nx^{n-2}\\ &\qquad+(1-3x)\sum_{n=1}^{+\infty}na_nx^{n-1}-\sum_{n=0}^{+\infty}a_nx^n\\ &=\sum_{\cancel{n=1}\ n=2}^{+\infty}n(n-1)a_nx^{n-1}-\sum_{\cancel{n=0}\ n=2}^{+\infty}n(n-1)a_nx^n\\ &\qquad+\sum_{n=1}^{+\infty}na_nx^{n-1}-3\sum_{\cancel{n=0}\ n=1}^{+\infty}na_nx^n\\ &\qquad-\sum_{n=0}^{+\infty}a_nx^n\\ &=\sum_{n=0}^{+\infty}n(n+1)a_{n+1}x^n-\sum_{n=0}^{+\infty}n(n-1)a_nx^n\\ &\qquad+\sum_{n=0}^{+\infty}(n+1)a_{n+1}x^n-3\sum_{n=0}^{+\infty}na_nx^n\\ &\qquad-\sum_{n=0}^{+\infty}a_nx^n\\ &=\sum_{n=0}^{+\infty}(n+1)^2(a_{n+1}-a_n)x^n\,.\end{aligned}$$

y est donc solution de l'équation différentielle si et seulement si $a_{n+1}=a_n$ pour

tout n ce qui donne :

$$y(x) = a_0 \sum_{n=0}^{+\infty} x^n = \frac{a_0}{1-x}\,,$$

cette série entière ayant un rayon de convergence égal à 1.

Réciproquement, la fonction $x \mapsto \dfrac{1}{1-x}$ est bien solution de l'équation (puisqu'on a trouvé un rayon de convergence strictement positif, ou bien par calcul direct), et on cherche alors les autres solutions sur $]0\,;1[$ par la méthode de variation de la constante.

En posant $y(x) = \dfrac{z(x)}{1-x}$, avec z de classe $\mathscr{C}^2$, on trouve que y est solution de l'équation proposée si et seulement si z est solution de l'équation :

$$xz'' + z' = 0 \quad (x \in]0\,;1[)\,.$$

(nous n'avons pas reproduit les calculs).
Il s'agit d'une équation linéaire du 1er ordre en z', homogène, dont la solution générale est :

$$z'(x) = \frac{\lambda}{x} \quad (\lambda \in \mathbb{R})\,,$$

d'où, pour $x \in]0\,;1[$:

$$z(x) = \lambda \ln x + \mu \quad (\lambda,\mu) \in \mathbb{R}^2\,.$$

Les solutions de l'équation proposée sur $]0\,;1[$ sont donc les fonctions de la forme

$$y\colon x \mapsto \frac{\lambda \ln x + \mu}{1-x} \quad (\lambda,\mu) \in \mathbb{R}^2\,.$$

Exercice 17.7

Sur $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$, on peut poser :

$$y(x) = \frac{z(x)}{x^2}\,,$$

avec z de classe $\mathscr{C}^2$. On a alors :

$$y = zx^{-2} \quad , \quad y' = z'x^{-2} - 2zx^{-3}$$
$$y'' = z''x^{-2} - 4z'x^{-3} + 6zx^{-4}\,.$$

En remplaçant dans l'équation, on trouve après un calcul facile que y est solution de l'équation différentielle (E) sur $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$ si et seulement si

$$z'' - z = 1\,,$$

ce qui donne :

$$z(x) = A\,\mathrm{ch}\,x + B\,\mathrm{sh}\,x - 1\,.$$

Les solutions de l'équation (E) sont donc leq applications de la forme :

$$y(x) = \begin{cases} \dfrac{A\,\mathrm{ch}\,x + B\,\mathrm{sh}\,x - 1}{x^2} & \text{si } x > 0 \\ \dfrac{C\,\mathrm{ch}\,x + D\,\mathrm{sh}\,x - 1}{x^2} & \text{si } x < 0 \end{cases}$$

avec $(A,B,C,D) \in \mathbb{R}^4$.

Puisque $\lim\limits_{x\to 0} \dfrac{\mathrm{ch}\,x - 1}{x^2}$ existe et vaut $\dfrac{1}{2}$, pour que y se prolonge en 0 il faut et il suffit que $A = C = 1$ et $B = D = 0$.

La seule solution possible sur $\mathbb{R}$ est donc :

$$f(x) = \begin{cases} \dfrac{\mathrm{ch}\,x - 1}{x^2} & \text{si } x \neq 0 \\ \dfrac{1}{2} & \text{si } x = 0, \end{cases}\,.$$

Cette application f est alors développable en série entière sur $\mathbb{R}$ (écrire le développement en série entière de la fonction ch), donc elle est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$ et ce sera bien la seule solution sur $\mathbb{R}$.

Exercice 17.8

1) D'après le cours, la solution cherchée est :

$$t \mapsto X(t) = \mathrm{e}^{tA} X_0\,.$$

Il faut donc calculer e^{tA}. Puisque $\lambda \mathrm{I}_n$ et B commutent, on a déjà :

$$\mathrm{e}^{tA} = \mathrm{e}^{\lambda t \mathrm{I}_n} \mathrm{e}^{tB} = \mathrm{e}^{\lambda t} \mathrm{e}^{tB}.$$

On a ensuite :

$$B^2 = U({}^tVU){}^tV = (\mathrm{tr}\,B)B$$

(en effet, pour $(i,j) \in [\![1\,;n]\!]^2$, $b_{ij} = u_i v_j$ donc ${}^tVU = \langle U \,|\, V\rangle = \sum\limits_{i=1}^{n} u_i v_i = \mathrm{tr}\,B$).
Donc par récurrence :

$$\forall\, k \in \mathbb{N}^*,\ B^k = (\mathrm{tr}\,B)^{k-1} B$$

puis :

$$\mathrm{e}^{tB} = \mathrm{I}_n + \left(\sum_{k=1}^{+\infty} \frac{(\mathrm{tr}\,B)^{k-1} t^k}{k!} \right) B.$$

Si $\mathrm{tr}\,B = 0$ alors $\mathrm{e}^{tB} = \mathrm{I}_n + tB$; sinon :

$$\mathrm{e}^{tB} = \mathrm{I}_n + \frac{\mathrm{e}^{t(\mathrm{tr}\,B)} - 1}{\mathrm{tr}\,B} B\,,$$

et, pour tout $t \in \mathbb{R}$:

$$X(t) = \mathrm{e}^{\lambda t} \mathrm{e}^{tB} X_0\,.$$

2) • B est de rang 1 puisque toutes ses colonnes sont proportionnelles à V et que $B \neq 0$ par hypothèse.
Puisque $n \geqslant 2$, $\dim \operatorname{Ker} B \geqslant 1$: il existe $X_0 \neq 0$ tel que $BX_0 = 0$. D'après le calcul fait à la question **1)**, on a alors $\mathrm{e}^{tB}X_0 = X_0$ d'où $X(t) = \mathrm{e}^{\lambda t}X_0$.
Une condition nécessaire pour que $\lim\limits_{t\to+\infty} X(t)$ existe est donc $\lambda \leqslant 0$.

• Supposons donc $\lambda \leqslant 0$ et X_0 tel que $BX_0 \neq 0$.

– Si $\operatorname{tr} B \neq 0$, on a :

$$X(t) = \mathrm{e}^{\lambda t}X_0 + \mathrm{e}^{\lambda t}\frac{\mathrm{e}^{t(\operatorname{tr} B)} - 1}{\operatorname{tr} B}BX_0.$$

Puisque $\lambda \leqslant 0$, il est clair que $\lim\limits_{t\to+\infty} X(t)$ existe si et seulement si $\lim\limits_{t\to+\infty} \mathrm{e}^{t(\lambda+\operatorname{tr} B)}$ existe, c'est-à-dire si $\lambda + \operatorname{tr} B \leqslant 0$.

– Si $\operatorname{tr} B = 0$, on a :

$$X(t) = \mathrm{e}^{\lambda t}(X_0 + tBX_0)\,,$$

donc $\lim\limits_{t\to+\infty} X(t)$ existe si et seulement si $\lambda < 0$

En conclusion, la condition nécessaire et suffisante cherchée est :

– $\lambda \leqslant \min(0, -\operatorname{tr} B)$ si $\operatorname{tr} B \neq 0$;

– $\lambda < 0$ si $\operatorname{tr} B = 0$.

Exercice 17.9

1) On peut déjà remarquer que la fonction nulle est évidemment solution. On pourra donc se contenter de ne chercher que les solutions non nulles.

• *Analyse*

On essaye ici de chercher tout ce qu'implique l'équation. Soit f une solution.

– Si l'on prend $y = 0$, on obtient $2f(x) = 2f(x)f(0)$ pour tout x réel. Puisqu'on a supposé que f n'était pas la fonction nulle, il existe un x tel que $f(x) \neq 0$, donc on peut en déduire : $f(0) = 2$.

– Si maintenant on choisit $x = 0$; on obtient, pour tout réel y,

$$f(y) + f(-y) = 2f(y)\,,$$

soit $f(-y) = f(y)$: f est nécessairement paire.

– Utilisons maintenant le fait que f est de classe $\mathscr{C}^2$. En dérivant deux fois l'équation par rapport à x, on obtient, pour tous x, y :

$$f''(x+y) + f''(x-y) = f''(x)f(y) \quad (1)$$

Et en la dérivant deux fois par rapport à y, on obtient :

$$f''(x+y) + f''(x-y) = 2f(x)f''(y) \quad (2)$$

En comparant (1) et (2), on a donc :

$$\forall\, x,\, y \in \mathbb{R},\ f''(x)f(y) = 2f(x)f''(y)\,.$$

En choisissant alors un y_0 tel que $f(y_0) \neq 0$, on obtient $f''(x) = \dfrac{f''(y_0)}{f(y_0)}f(x)$, donc il existe un réel λ tel que :

$$\forall\, x \in \mathbb{R},\ f''(x) = \lambda f(x)\,.$$

• *Synthèse*

Les solutions autres que la fonction nulle sont donc à chercher parmi les fonctions f telles que :

$$f(0) = 2 \quad ; \quad f \text{ est paire} \quad ;$$
$$\exists\, \lambda \in \mathbb{R} \text{ tel que } f'' = \lambda f\,.$$

Pour l'équation différentielle, il y a trois possibilités.

– Si $\lambda = 0$: f est nécessairement de la forme $x \mapsto ax + b$, et compte tenu des deux contraintes supplémentaires, seule fa fonction constante égale à 2 convient.

– Si $\lambda > 0$: f est nécessairement de la forme $x \mapsto a\operatorname{ch}(\sqrt{\lambda}x) + b\operatorname{sh}(\sqrt{\lambda}x)$. f étant paire, le coefficient b est nul, et puisque $f(0) = 2$, on a nécessairement $f(x) = 2\operatorname{ch}(\sqrt{\lambda}x)$.

– Si $\lambda < 0$: f est nécessairement de la forme $x \mapsto a\cos(\sqrt{-\lambda}x) + b\sin(\sqrt{-\lambda}x)$. f étant paire, le coefficient b est nul, et puisque $f(0) = 2$, on a nécessairement $f(x) = 2\cos(\sqrt{-\lambda}x)$.

Il est ensuite indispensable de vérifier que les fonctions obtenues conviennent, ce qui découle immédiatement des formules de trigonométrie habituelles.

En conclusion, les solutions à l'équation proposée sont : les deux applications constantes égales à 0 et 2, les applications de la forme $x \mapsto 2\cos ax$ et les applications de la forme $x \mapsto 2\operatorname{ch} ax$ avec $a \in \mathbb{R}$.

2) Supposons maintenant f seulement continue et solution de l'équation. Puisque l'on exclut la solution nulle, il existe un réel x_0 tel que $x_0 \neq 0$.
f étant continue, il existe un intervalle fermé $I = [a\,;b]$ de centre x_0 sur lequel f ne s'annule pas.
On aura alors pour tout x, en intégrant la relation proposée par rapport à y :

$$\int_a^b \big(f(x+y)+f(x-y)\big)\,\mathrm{d}y = 2f(x)\int_a^b f(y)\,\mathrm{d}y = Cf(x)\,,$$

en notant $C = 2\displaystyle\int_a^b f(y)\,\mathrm{d}y$. f étant continue et ne s'annulant pas sur $[a\,;b]$, on a $C \neq 0$ d'où pour tout $x \in \mathbb{R}$:

$$f(x) = \frac{1}{C}\int_a^b \big(f(x+y)+f(x-y)\big)\,\mathrm{d}y\,,$$

donc en notant F une primitive de f :

$$f(x) = \frac{1}{C}\big(F(x+b)-F(x+a) - F(x-b)+F(x-a)\big),$$

ce qui implique que f est de classe $\mathscr{C}^1$. Mais F est alors de classe $\mathscr{C}^2$, et f aussi (en fait, on montrerait ainsi que f est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$).
On est ainsi ramené au cas précédent.

Exercice 17.10

1) • *Analyse*
Supposons qu'il existe deux solution y_1 et y_2 vérifiant $y_2 = xy_1$ sur I . Alors

$$y_2' = xy_1' + y_1 \quad \text{et} \quad y_2'' = xy_1'' + 2y_1'$$

d'où pour tout $x \in I$:

$$\begin{aligned} 0 &= y_2'' + a(x)y_2' + b(x)y_2 \\ &= xy_1'' + 2y_1' + a(x)(xy_1' + y_1) + b(x)xy_1 \\ &= x\underbrace{(y_1'' + a(x)y_1' + b(x)y_1)}_{=0} + 2y_1' + a(x)y_1 \\ &= 2y_1' + a(x)y_1\,. \end{aligned}$$

En dérivant cette relation, on obtient :

$$2y_1'' + a(x)y_1' + a'(x)y_1 = 0$$

d'où :

$$2(-a(x)y_1' - b(x)y_1) + a(x)y_1' + a'(x)y_1 = 0$$

et enfin :

$$-a(x)y_1' + (a'(x) - 2b(x))y_1 = 0\,.$$

Compte tenu de la relation $2y_1' + a(x)y_1 = 0$ trouvée au début, qui implique $-a(x)y_1' = \frac{1}{2}a(x)^2y_1$, on en tire :

$$\big(a(x)^2 + 2a'(x) - 4b(x)\big)\,y_1 = 0\,.$$

S'il existe $x_0 \in I$ tel que $y_1(x_0) = 0$, la relation $2y_1' + a(x)y_1 = 0$ donne alors $y_1'(x_0) = 0$ et, d'après le théorème de Cauchy-Lipschitz, y_1 est identiquement nulle, ce qui est exclu.
Donc y_1 ne peut s'annuler et on obtient la condition nécessaire :

$$\forall\, x \in I,\ a(x)^2 + 2a'(x) - 4b(x) = 0\,.$$

• *Synthèse*
Réciproquement, supposons cette condition vérifiée, et soit y_1 une solution de l'équation linéaire du 1er ordre :

$$2y_1' + a(x)y_1 = 0\,.$$

Alors y_1' est dérivable ; en dérivant on a :

$$2y_1'' + a(x)y_1' + a'(x)y_1 = 0$$

donc

$$\begin{aligned} &y_1'' + a(x)y_1' + b(x)y_1 \\ &\quad = \frac{a(x)y_1'}{2} - \frac{a'(x)y_1}{2} + b(x)y_1 \\ &\quad = -\frac{a(x)^2y_1}{4} - \frac{a'(x)y_1}{2} + b(x)y_1 = 0 \end{aligned}$$

(d'après la relation vérifiée par $a(x)$), ce qui montre que y_1 est solution de l'équation différentielle proposée.
En posant alors $y_2 = xy_1$ on a :

$$\begin{aligned} &y_2'' + a(x)y_2' + b(x)y_2 \\ &\quad = xy_1'' + 2y_1' + a(x)(xy_1' + y_1) + b(x)xy_1 \\ &\quad = x\underbrace{(y_1'' + a(x)y_1' + b(x)y_1)}_{=0} + 2y_1' + a(x)y_1 \\ &\quad = 2y_1' + a(x)y_1 = 0\,, \end{aligned}$$

ce qui montre que y_2 est aussi solution de l'équation.

• En conclusion, la condition nécessaire et suffisante cherchée est :

$$\forall\, x \in I,\ a(x)^2 + 2a'(x) - 4b(x) = 0\,.$$

2) On vérifie déjà que la condition ci-dessus est vérifiée pour l'équation proposée (ici, $a(x) = 2x$ et $b(x) = 1 + x^2$).
On utilise ensuite les résultats du calcul précédent, en considérant une solution y_1

de l'équation homogène $2y_1' + a(x)y_1 = 0$, qui s'écrit ici :

$$y_1' + xy_1 = 0\,.$$

On peut prendre $y_1 : x \mapsto \mathrm{e}^{-\frac{x^2}{2}}$. Alors y_1 et xy_1 sont deux solutions linéairement indépendantes, donc l'ensemble des solutions sur $\mathbb{R}$ de l'équation différentielle

$$y'' + 2xy' + (1+x^2)y = 0$$

est l'ensemble des applications de la forme :

$$x \mapsto (ax+b)\mathrm{e}^{-\frac{x^2}{2}} \quad (a,b) \in \mathbb{R}^2\,.$$

Exercice 17.11

1) Soit y une solution sur $\mathbb{R}$ de l'équation :

$$y'' + p(x)y = 0\,.$$

Si, par l'absurde, y ne s'annule pas sur $\mathbb{R}$, elle y garde un signe constant d'après le théorème des valeurs intermédiaires. Quitte à changer y en $-y$ (qui est aussi solution), on peut supposer y à valeurs strictement positives.
Puisque $y'' = -py$, y'' est négative donc y est concave. Sa courbe est donc située en-dessous de ses tangentes, ce qui s'écrit, pour tous x_0, $x \in \mathbb{R}$:

$$y(x) \leqslant y(x_0) + (x-x_0)y'(x_0)\,.$$

S'il existait un x_0 tel que $y'(x_0) < 0$, on en déduirait $\lim\limits_{x\to+\infty} y(x) = -\infty$, ce qui contredit l'hypothèse y positive.

S'il existait un x_0 tel que $y'(x_0) > 0$, on en déduirait $\lim\limits_{x\to-\infty} y(x) = -\infty$, ce qui contredit encore la positivité de y

On a donc forcément $y' = 0$ donc y est constante. Sachant que y est solution de l'équation $y'' + py = 0$ et que p n'est pas la fonction nulle, cela implique que y est la fonction nulle, ce qui contredit encore l'hypothèse.

On a donc démontré par l'absurde que y s'annule au moins une fois.

2) Soit y une solution sur $\mathbb{R}$ de de l'équation :

$$y'' - p(x)y = 0\,.$$

• Si y s'annulait en deux points a et b avec $a < b$, on aurait, en multipliant l'équation par y et en intégrant entre a et b : :

$$\begin{aligned}\int_a^b p(x)y^2(x)\,\mathrm{d}x &= \int_a^b y(x)y''(x)\,\mathrm{d}x\\ &= \big[y'(x)y(x)\big]_a^b - \int_a^b y'^2(x)\,\mathrm{d}x\\ &= -\int_a^b y'^2(x)\,\mathrm{d}x\,.\end{aligned}$$

Puisqu'on a supposé $p \geqslant 0$, l'intégrale $\int_a^b py^2$ est positive donc on doit avoir :

$$\int_a^b p(x)y^2(x)\,\mathrm{d}x = -\int_a^b y'^2(x)\,\mathrm{d}x = 0\,.$$

y'^2 étant une fonction continue positive, elle est nulle sur $[a\,;b]$ donc y est constante sur $[a\,;b]$ et puisque $y(a) = y(b) = 0$, y est nulle sur $[a\,;b]$. Mais on aurait alors $y'(a) = 0$, et d'après le théorème de Cauchy-Lipschitz, y serait la solution nulle, ce qui est exclu.

On a donc prouvé que y s'annule au plus une fois.

• Si y' s'annulait en deux points $a < b$, le calcul ci-dessus donne le même résultat puisque $\big[y'(x)y(x)\big]_a^b = 0$. On conclut donc de la même façon.

Exercice 17.12

1) Cette question a été traitée en **[SF17.13]** page 630.

2) a) Notons W le wronskien de y_1 et y_2 :

$$W = y_1y_2' - y_1'y_2\,.$$

Puisque y_1 et y_2 sont indépendantes, on sait que W ne s'annule pas.
Or, s'il existait t_0 tel que $y_1(t_0) = y_2(t_0) = 0$, on aurait $W(t_0) = 0$: contradiction.

b) Soient $a < b$ deux zéros consécutifs de y_1. Puisque y_1 ne s'annule pas sur $]a\,;b[$, la fonction $\varphi = \dfrac{y_2}{y_1}$ est de classe $\mathscr{C}^1$ sur cet intervalle et :

$$\forall\, t \in\,]a\,;b[,\ \varphi'(t) = \frac{W(t)}{y_1(t)^2}\,.$$

Puisque $W(t)$ garde un signe constant sur $]a\,;b[$, il en est de même de φ', et φ est strictement monotone.
Puisque $y_2(a)$ et $y_2(b)$ sont non nuls, la limite de φ en a^+ et en b^- est $\pm\infty$; φ étant strictement monotone, ces limites sont de signes différents. φ réalise donc une bijection de $]a\,;b[$ sur $\mathbb{R}$, et par suite

s'annule en un point c et un seul, et il en est de même de y_2.

Exercice 18.1

D'après les théorèmes généraux, f est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$, et il suffit donc d'étudier sa continuité en (0,0), ce qui revient à étudier sa limite en ce point.
Pour $(x,y) \neq (0,0)$, posons

$$r = \sqrt{|x|^c + |y|^d}\,,$$

de sorte que :

$$\left(\frac{|x|^{c/2}}{r}\right)^2 + \left(\frac{|y|^{d/2}}{r}\right)^2 = 1\,.$$

On peut donc poser :

$$|x|^{c/2} = r\cos\theta \quad \text{et} \quad |y|^{d/2} = r\sin\theta\,.$$

On aura alors :

$$\begin{aligned} f(x,y) &= \frac{(r\cos\theta)^{\frac{2a}{c}} \cdot (r\sin\theta)^{\frac{2b}{d}}}{r^2} \\ &= r^{\frac{2a}{c}+\frac{2b}{d}-2}(\cos\theta)^{\frac{2a}{c}}(\sin\theta)^{\frac{2b}{d}}\,. \end{aligned}$$

Ainsi, $|f(x,y)| \leqslant r^{2\left(\frac{a}{c}+\frac{b}{d}-1\right)}$, donc si $\frac{a}{c}+\frac{b}{d} > 1$, $\lim\limits_{(x,y)\to(0,0)} f(x,y) = f(0,0) = 0$, et f est bien continue en (0,0).

Si $\frac{a}{c}+\frac{b}{d} \leqslant 1$, on a alors :

$$\lim_{x\to 0} f\left(|x|^{\frac{1}{c}}, |x|^{\frac{1}{d}}\right) = \lim_{x\to 0} \frac{|x|^{\frac{a}{c}+\frac{b}{d}}}{2\,|x|} \neq 0\,,$$

ce qui montre que f n'est pas continue en (0,0).

Exercice 18.2

Les théorèmes généraux permettent d'affirmer que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$. Le seul problème pour les trois questions est donc en (0,0).

1) En posant $r = \sqrt{x^2+y^2}$ on a :

$$|f(x,y)| \leqslant r^{2a}$$

donc si $a > 0$, $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0$, ce qui montre que f est aussi continue à l'origine.

Lorsque $a \leqslant 0$, considérons, pour $k \in \mathbb{N}^*$,

$$f\left(\frac{2}{k\pi},0\right) = \left(\frac{4}{k^2\pi^2}\right)^a \sin\left(\frac{k\pi}{2}\right)\,.$$

Les $\left(\frac{2}{k\pi},0\right)$ pour $k \in \mathbb{N}^*$ forment donc une suite de points qui tend vers (0,0) quand $k \to +\infty$, mais tels que leur image par f ne tend pas vers $f(0,0)$, ce qui montre que f n'est pas continue en (0,0).

2) • Cherchons d'abord les dérivées partielles en (0,0). Pour $x \neq 0$ on a :

$$\frac{f(x,0)-f(0,0)}{x} = x^{2a-1}\sin\frac{1}{|x|}\,,$$

et la limite de cette expression quand $x \to 0$ existe dans $\mathbb{R}$ si et seulement si $a > \frac{1}{2}$.

Ainsi, $\frac{\partial f}{\partial x}(0,0)$ existe si et seulement si $a > \frac{1}{2}$, et dans ce cas $\frac{\partial f}{\partial x}(0,0) = 0$.
Par symétrie, il en est de même de la dérivée partielle par rapport à y.

• Une condition nécessaire pour que f soit différentiable en (0,0) est donc $a > \frac{1}{2}$. Et dans ce cas, puisque :

$$|f(x,y) - f(0,0)| \leqslant r^{2a}\,,$$

on a $f(x,y) \underset{r\to 0}{=} f(0,0) + o(r)$, ce qui montre par définition que f est différentiable en (0,0) et que sa différentielle en ce point est nulle.

3) • Pour $(x,y) \neq (0,0)$:

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= 2ax(x^2+y^2)^{a-1}\sin\frac{1}{\sqrt{x^2+y^2}} \\ &\quad - x(x^2+y^2)^{a-3/2}\cos\frac{1}{\sqrt{x^2+y^2}} \end{aligned}$$

En posant encore $r = \sqrt{x^2+y^2}$:

$$\begin{aligned} \left|\frac{\partial f}{\partial x}(x,y)\right| &\leqslant 2ar.r^{2(a-1)} + r.r^{2(a-3/2)} \\ &\leqslant 2ar^{2a-1} + r^{2a-2}\,, \end{aligned}$$

donc si $a > 1$,

$$\lim_{(x,y)\to(0,0)} \frac{\partial f}{\partial x}(x,y) = 0 = \frac{\partial f}{\partial x}(0,0)\,.$$

Dans ce cas la dérivée partielle $\frac{\partial f}{\partial x}$ est continue en (0,0), et il en est de même par symétrie pour $\frac{\partial f}{\partial y}$. f est donc de classe $\mathscr{C}^1$.

• Enfin, dans le cas $\frac{1}{2} < a \leqslant 1$, la même suite de points que celle de la 1re question permet de montrer que $\frac{\partial f}{\partial x}$ n'est pas continue à l'origine, donc f n'est pas de classe $\mathscr{C}^1$.

Exercice 18.3

Pour $x = (x_1, \ldots, x_n)$ on a :

$$f(x) = \frac{1}{\sqrt{\sum_{i=1}^{n} x_i^2}} \cdot$$

L'application $x \mapsto \sum_{i=1}^{n} x_i^2$ étant de classe $\mathscr{C}^1$ de $\mathbb{R}^n \backslash \{(0,0)\}$ dans $\mathbb{R}_+^*$, la fonction $\sqrt{\ }$ étant de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$, et la fonction $y \mapsto \frac{1}{y}$ étant de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$, par composition, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^n \setminus \{(0,0)\}$.

Pour tout $i \in [\![1\,;n]\!]$ on aura alors :

$$\frac{\partial f}{\partial x_i}(x) = -x_i \left(\sum_{i=1}^{n} x_i^2 \right)^{-3/2} = -\frac{x_i}{\|x\|^3} \cdot$$

Le gradient ∇f de f en x étant le vecteur de coordonnées les $\frac{\partial f}{\partial x_i}$, on aura donc :

$$\forall\, x \neq 0,\ \nabla f(x) = -\frac{x}{\|x\|^3} \cdot$$

Exercice 18.4

Pour simplifier l'écriture, dans toute la suite de l'exercice on notera $U = \mathbb{R}_+^* \times \mathbb{R}$. Il s'agit bien d'un ouvert de $\mathbb{R}^2$.

1) Montrons que φ est bijective de U sur U et que φ et φ^{-1} sont de classe $\mathscr{C}^1$.

• Pour montrer que φ est bijective de U sur U, il suffit de revenir à la définition, c'est-à-dire montrer que :

$$\forall\,(u,v) \in U,\ \exists !\,(x,y) \in U \quad \text{tel que } (u,v) = \varphi(x,y)\,.$$

Soient donc $u,\ v \in \mathbb{R}$ tels que $u > 0$. L'équation $(u,v) = \varphi(x,y)$ équivaut à

$$\begin{cases} u = x \\ v = \dfrac{y}{x} \end{cases} \quad \text{soit} \quad \begin{cases} x = u \\ y = uv \end{cases} \cdot$$

Cette équation a donc bien une et une seule solution $(x,y) \in U$ car $x > 0$, ce qui prouve que φ est bijective de U sur U. On a également démontré par ce calcul que φ^{-1} est l'application qui à (u,v) associe (u,uv).

• Puisque x ne s'annule pas :

$$\forall\,(x,y) \in U,\ \frac{\partial \varphi}{\partial x}(x,y) = \left(1, -\frac{y}{x^2}\right) \quad \text{et} \quad \frac{\partial \varphi}{\partial y}(x,y) = \left(0, \frac{1}{x}\right) \cdot$$

Les dérivées partielles premières de φ sont donc bien des applications continues sur U, ce qui prouve que φ est de classe $\mathscr{C}^1$ sur U.

• De même, $\varphi^{-1} : (u,v) \mapsto (u,uv)$ est de classe $\mathscr{C}^1$ sur U.

2) Pour résoudre l'équation aux dérivées partielles proposée, on utilise le changement de variables suggéré par l'énoncé, en posant, pour tout $(x,y) \in U$:

$$f(x,y) = f \circ \varphi^{-1}(u,v) = F(u,v)\,.$$

On peut alors exprimer les dérivées partielles de f par rapport à x et y en fonction de celles de F par rapport à u et v à l'aide des formules pour les fonctions composées :

$$\begin{cases} \dfrac{\partial f}{\partial x} = \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial x} \\ \dfrac{\partial f}{\partial y} = \dfrac{\partial F}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial F}{\partial v}\dfrac{\partial v}{\partial y} \end{cases} \cdot$$

Puisque ici $u = x$ et $v = \dfrac{y}{x}$, le système précédent devient simplement :

$$\begin{cases} \dfrac{\partial f}{\partial x} = \dfrac{\partial F}{\partial u} - \dfrac{y}{x^2}\dfrac{\partial F}{\partial v} \\ \dfrac{\partial f}{\partial y} = \dfrac{1}{x}\dfrac{\partial F}{\partial v} \end{cases} \quad (*)$$

donc $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = x\dfrac{\partial F}{\partial u}$.

Puisque $x > 0$ sur U, l'équation aux dérivées partielles $x\dfrac{\partial f}{\partial x} + y\dfrac{\partial f}{\partial y} = 0$ est donc équivalente à :

$$\frac{\partial F}{\partial u} = 0\,.$$

L'ensemble des solutions de cette dernière équation est l'ensemble des fonctions F telles que $F(u,v) = g(v)$ pour tout $(u,v) \in U$, avec g de classe $\mathscr{C}^1$ sur $\mathbb{R}$, quelconque.

En conclusion, l'ensemble des solutions sur U de l'équation :

$$x\frac{\partial f}{\partial x}(x,y) + y\frac{\partial f}{\partial y}(x,y) = 0$$

est l'ensemble des fonctions f de la forme :

$$f(x,y) = g\left(\frac{y}{x}\right) \text{ avec } g \in \mathscr{C}^1(\mathbb{R},\mathbb{R})\,.$$

✓ Compte tenu du résultat, on se rend compte qu'un changement en coordonnées polaires eût aussi été possible. Vous pouvez essayer de vous entraîner avec le changement de variables

$x = r\cos\theta$, $y = r\sin\theta$ avec $(r,\theta) \in \mathbb{R}_+^* \times \left]-\frac{\pi}{2}\,;\frac{\pi}{2}\right[$.

3) On considère maintenant l'équation aux dérivées partielles :

$$x^2\,\frac{\partial^2 f}{\partial x^2} + 2xy\,\frac{\partial^2 f}{\partial x \partial y} + y^2\,\frac{\partial^2 f}{\partial y^2} = 0\,,$$

pour laquelle on utilisera le même changement de variables.
Il est facile de vérifier que φ et φ^{-1} sont de classe $\mathscr{C}^2$ sur U (toutes les dérivées partielles d'ordre 2 existent et sont continues), de sorte que la fonction F sera elle aussi de classe $\mathscr{C}^2$ si f l'est.

On part alors des relations $(*)$ obtenues précédemment, et on applique à $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ le même principe que celui que l'on a appliqué à f.
Cela donne les affreux calculs suivants (où l'on a utilisé le théorème de Schwarz pour calculer les dérivées partielles d'ordre deux dans un ordre de dérivation indifférent) :

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u} - \frac{y}{x^2}\,\frac{\partial F}{\partial v}\right)\\
&= \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial u}\right) - \frac{y}{x^2}\,\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial v}\right) + \frac{2y}{x^3}\,\frac{\partial F}{\partial v}\\
&= \frac{\partial^2 F}{\partial u^2}\,\frac{\partial u}{\partial x} + \frac{\partial^2 F}{\partial u \partial v}\,\frac{\partial v}{\partial x}\\
&\qquad - \frac{y}{x^2}\left(\frac{\partial^2 F}{\partial u \partial v}\,\frac{\partial u}{\partial x} + \frac{\partial^2 F}{\partial v^2}\,\frac{\partial v}{\partial x}\right)\\
&\qquad + \frac{2y}{x^3}\,\frac{\partial F}{\partial v}\\
&= \frac{\partial^2 F}{\partial u^2} - 2\frac{y}{x^2}\,\frac{\partial^2 F}{\partial u \partial v}\\
&\qquad + \frac{y^2}{x^4}\,\frac{\partial^2 F}{\partial v^2} + \frac{2y}{x^3}\,\frac{\partial F}{\partial v} \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{1}{x}\,\frac{\partial F}{\partial v}\right)\\
&= -\frac{1}{x^2}\,\frac{\partial F}{\partial v} + \frac{1}{x}\,\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial v}\right)\\
&= -\frac{1}{x^2}\,\frac{\partial F}{\partial v} + \frac{1}{x}\left(\frac{\partial^2 F}{\partial u \partial v}\,\frac{\partial u}{\partial x} + \frac{\partial^2 F}{\partial v^2}\,\frac{\partial v}{\partial x}\right)\\
&= -\frac{1}{x^2}\,\frac{\partial F}{\partial v} + \frac{1}{x}\,\frac{\partial^2 F}{\partial u \partial v} - \frac{y}{x^3}\,\frac{\partial^2 F}{\partial v^2} \qquad (2)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{1}{x}\,\frac{\partial F}{\partial v}\right) = \frac{1}{x}\,\frac{\partial}{\partial y}\left(\frac{\partial F}{\partial v}\right)\\
&= \frac{1}{x}\left(\frac{\partial^2 F}{\partial u \partial v}\,\frac{\partial u}{\partial y} + \frac{\partial^2 F}{\partial v^2}\,\frac{\partial v}{\partial y}\right)\\
&= \frac{1}{x^2}\,\frac{\partial^2 F}{\partial v^2} \qquad (3)
\end{aligned}$$

On a donc finalement, en faisant $x^2 \times (1) + 2xy \times (2) + y^2 \times (3)$:

$$x^2\,\frac{\partial^2 f}{\partial x^2} + 2xy\,\frac{\partial^2 f}{\partial x \partial y} + y^2\,\frac{\partial^2 f}{\partial y^2} = x^2\,\frac{\partial^2 F}{\partial u^2}\,.$$

Puisque $x \neq 0$ sur U, l'équation différentielle de l'énoncé équivaut donc à :

$$\frac{\partial^2 F}{\partial u^2}(u,v) = 0\,.$$

Cela s'intègre successivement en $\frac{\partial F}{\partial u} = g(v)$ puis $F(u,v) = ug(v) + h(v)$, où g et h sont des fonctions de classe $\mathscr{C}^2$ sur $\mathbb{R}$.

En conclusion, l'ensemble des solutions sur U de l'équation (E_2) est l'ensemble des fonctions de la forme :

$$f(x,y) = xg\left(\frac{y}{x}\right) + h\left(\frac{y}{x}\right)\,,$$

où $g,\, h \in \mathscr{C}^2(\mathbb{R},\mathbb{R})$.

Exercice 18.5

1) f étant symétrique, il existe une base orthonormale formée de vecteurs propres de f (c'est le théorème spectral).
Soit donc $\mathscr{B} = (e_1,\dots,e_n)$ une telle base, chaque e_i étant associé à une valeur propre λ_i, strictement positive par hypothèse.

Si $x \in \mathbb{R}^n \setminus \{0\}$ a pour coordonnées $(x_1,\dots,x_n)$ dans cette base, on a $x = \sum_{i=1}^{n} x_i e_i$, d'où :

$$f(x) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} \lambda_i x_i e_i\,,$$

et puisque la base est orthonormale, on en déduit :

$$\langle f(x)\,|\,x\rangle = \sum_{i=1}^{n} \lambda_i x_i^2$$

donc, puisque les x_i ne sont pas tous nuls et les $\lambda_i > 0$:

$$\forall\, x \in \mathbb{R}^n \setminus \{0\},\ \langle f(x)\,|\,x\rangle > 0.$$

2) Soit $a \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$. Alors, en utilisant la linéarité de f et la bilinéarité du produit scalaire :

$$g(a+h) - g(a) = \frac{1}{2}\langle f(a)\,|\,h\rangle + \frac{1}{2}\langle f(h)\,|\,a\rangle - \langle u\,|\,h\rangle + \frac{1}{2}\langle f(h)\,|\,h\rangle \ .$$

Notons, pour tout $h \in \mathbb{R}^n$,

$$\ell(h) = \frac{1}{2}\langle f(a)\,|\,h\rangle + \frac{1}{2}\langle f(h)\,|\,a\rangle - \langle u\,|\,h\rangle \ .$$

f étant symétrique, on a

$$\langle f(a)\,|\,h\rangle = \langle a\,|\,f(h)\rangle \ ,$$

de sorte que

$$\ell(h) = \langle f(a) - u\,|\,h\rangle \ .$$

Il est clair que ℓ est un endomorphisme de $\mathbb{R}^n$, et l'on a :

$$\forall\, a \in \mathbb{R}^n,\ \forall\, h \in \mathbb{R}^n,\quad g(a+h) - g(a) = \ell(h) + \frac{1}{2}\langle f(h)\,|\,h\rangle \quad (*)$$

L'application linéaire f est continue puisque $\mathbb{R}^n$ est un $\mathbb{R}$-espace vectoriel de dimension finie, donc on sait qu'il existe une constante réelle k telle que :

$$\forall\, h \in \mathbb{R}^n,\ \|f(h)\| \leqslant k\,\|h\| \ .$$

L'inégalité de Cauchy-Schwarz donne alors :

$$|\langle f(h)\,|\,h\rangle| \leqslant \|f(h)\| \cdot \|h\| \leqslant k \cdot \|h\|^2 \ ,$$

ce qui prouve que

$$|\langle f(h)\,|\,h\rangle| = O\big(\|h\|^2\big) = o(\|h\|) \ ,$$

et la relation $(*)$ est alors exactement la définition de la différentiabilité de g en a.

En conclusion, g est différentiable en tout point a de $\mathbb{R}^n$, et pour tout $h \in \mathbb{R}^n$,

$$\mathrm{d}g_a(h) = \langle f(a) - u\,|\,h\rangle \ .$$

3) • a est un point critique de g si et seulement si $dg_a = 0$, c'est-à-dire si

$$\forall\, h \in \mathbb{R}^n \ ,\ \langle f(a) - u\,|\,h\rangle = 0 \, ,$$

ce qui équivaut à $f(a) = u$. Or f est de déterminant strictement positif (produit des valeurs propres strictement positives), donc c'est un automorphisme de $\mathbb{R}^n$.
Par conséquent, g admet un et un seul point critique : $a = f^{-1}(u)$.

• Il suffit de reprendre les calculs faits plus haut (relation $(*)$) :

$$\forall\, a \in \mathbb{R}^n,\ \forall\, h \in \mathbb{R}^n,\quad g(a+h) - g(a) = \mathrm{d}g_a(h) + \frac{1}{2}\langle f(h)\,|\,h\rangle \ .$$

Pour $a = f^{-1}(u)$, on a alors $\mathrm{d}g_a(h) = 0$ donc $g(a+h) - g(a) = \frac{1}{2}\langle f(h)\,|\,h\rangle \geqslant 0$ (avec inégalité stricte si $h \neq 0$).
Par conséquent, g admet en a un minimum global strict.

Exercice 18.6

On peut déjà remarquer que le domaine D est l'intérieur (bord compris) d'une ellipse (qui est facile à représenter). De plus, D est un *fermé* ; en effet, D n'est rien d'autre que l'image réciproque par l'application $(x,y) \mapsto x^2 + 2y^2$, qui est continue de $\mathbb{R}^2$ dans $\mathbb{R}$, de l'intervalle $]-\infty\,;8]$, qui est une partie fermée de $\mathbb{R}$. Et on sait que l'image réciproque d'un fermé par une application continue est un fermé.

De plus, D est évidemment borné.

D'après les théorèmes usuels, on peut affirmer que f est continue sur $\mathbb{R}^2$. En effet, les applications linéaires coordonnées $(x,y) \mapsto x$ et $(x,y) \mapsto y$ le sont (résultat du cours), et il s'agit ensuite de sommes, produits et composées de fonctions continues d'une variable réelle.

D'après un théorème fondamental du cours, f étant continue sur la partie *fermée bornée* D y est bornée et atteint ses bornes.

On cherche d'abord les extrema de f sur l'intérieur $\overset{\circ}{D}$ de D. Malheureusement, f n'est pas de classe $\mathscr{C}^1$ sur l'intérieur de D (en effet, la fonction réelle $x \mapsto \sqrt{x}$ n'est pas dérivable en 0) ; par suite, les dérivées partielles de f (calculées ci-après) sont seulement définies sur $\mathbb{R}^2 \setminus \{(0,0)\}$. On examine donc 2 cas.

- En $(0,0)$, on a $f(0,0) = 0$ et, puisque f est à valeurs positives, elle admet en ce point un minimum global.
- Sur $\overset{\circ}{D}$ privé de $(0,0)$, qui est un ouvert, f est de classe $\mathscr{C}^1$; ses extrema locaux sur cet ensemble sont donc des points critiques (c'est-à-dire des points où le gradient de f s'annule).

On calcule donc, pour $(x,y) \neq (0,0)$:

$$\frac{\partial f}{\partial x}(x,y) = \frac{x}{\sqrt{x^2+y^2}} + 2x\,,$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{y}{\sqrt{x^2+y^2}}\,,$$

et les relations $\frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = 0$ conduisent facilement à $x = y = 0$, qui est un cas exclu.

En conclusion, f ne peut pas admettre d'extremum local sur $\overset{\circ}{D}\backslash\{(0,0)\}$, et son seul extremum sur $\overset{\circ}{D}$ est un minimum global, atteint en $(0,0)$.

Il reste donc à examiner les extrema de f sur le bord de D, c'est-à-dire sur l'ellipse d'équation $x^2 + 2y^2 = 8$. En un point (x,y) du bord, puisque $y^2 = 4 - \frac{x^2}{2}$, on a $f(x,y) = \sqrt{\frac{x^2}{2} + 4} + x^2$. On est donc ramené à chercher les extrema de la fonction :

$$\varphi\colon x \mapsto \sqrt{\frac{x^2}{2} + 4} + x^2$$

lorsque x décrit l'intervalle $\left[-2\sqrt{2}\,;2\sqrt{2}\right]$ (car $x^2 \leqslant 8$).

Ici, nul besoin d'étude compliquée ! φ est évidemment paire, et elle est croissante sur $[0,2\sqrt{2}]$ comme composée et somme de fonctions croissantes sur cet intervalle.
Il en résulte que φ admet un minimum pour $x = 0$, et un maximum pour $x = \pm 2\sqrt{2}$.

Les extrema de la restriction de f au bord de D sont donc atteints pour $x = 0$, c'est-à-dire aux points $(0,2)$ et $(0,-2)$, et pour $x = \pm 2\sqrt{2}$, c'est-à-dire aux points $(2\sqrt{2},0)$ et $(-2\sqrt{2},0)$ (ce sont les 4 sommets de l'ellipse).

Les extrema de f sur D autres que $(0,0)$ sont donc à chercher *parmi* ces 4 points. Examinons le comportement de f au voisinage de ces points :

- En $(0,\pm 2)$, f prend la valeur 2, qui est son minimum *sur le bord*. Mais, en des points de D au voisinage de ce point, par exemple des points de la forme $(0,\pm a)$ avec a proche de 2 et inférieur à 2, on a $f(0,\pm a) = a < 2$. Donc f prend au voisinage de $(0,\pm 2)$ à la fois des valeurs supérieures (sur le bord de D) et des valeurs inférieures (sur l'axe Oy). Ce n'est donc pas un extremum local.
- En $(\pm 2\sqrt{2},0)$, f prend la même valeur (peu importe laquelle !) qui est son maximum sur le bord. Or, on sait que f admet un maximum sur D, et que ce maximum ne peut être atteint que sur le bord. C'est donc nécessairement en ces points.

✓ Il est important de retenir le raisonnement précédent, qui s'applique chaque fois que l'on veut rechercher des extrema d'une fonction de plusieurs variables sur une partie fermée bornée.

Pour conclure : f admet sur D un minimum global en $(0,0)$ et un maximum global en $(\pm 2\sqrt{2},0)$.

Exercice 18.7

L'équation de (S) s'écrit $f(x,y,z) = 0$ avec $f(x,y,z) = x^2 + y^2 + 2z^2 - 1 = 0$.
f est de classe $\mathscr{C}^1$ et pour tout $(x,y,z) \in \mathbb{R}^3$, on a :

$$\nabla f(x,y,z) = (2x,2y,4z)\ .$$

Puisque $(0,0,0)$ n'appartient pas à (S), le gradient de f en un point de (S) n'est jamais nul. Le plan tangent en un point $M_0(x_0,y_0,z_0)$ de (S) aura donc pour équation :

$$2x_0(x-x_0) + 2y_0(y-y_0) + 4z_0(z-z_0) = 0\,,$$

soit, compte tenu de la relation $x_0^2 + y_0^2 + 2z_0^2 = 1$:

$$x_0x + y_0y + 2z_0z - 1 = 0\,.$$

La droite d'équation

$$x = \frac{y}{3} = -\frac{z}{2}$$

a pour vecteur directeur $(1,3,-2)$. Elle est donc orthogonale au plan précédent si les deux vecteurs $(1,3,-2)$ et $(x_0,y_0,2z_0)$ sont colinéaires, soit $y_0 = 3x_0$ et $z_0 = -x_0$.

Compte tenu de la relation $x_0^2 + y_0^2 + 2z_0^2 = 1$, on trouve $x_0^2 = \frac{1}{12}$, il y a donc deux plans qui conviennent.

Table des matières

Présentation de la collection 3

Comment utiliser ce livre ? 4

Quelques conseils pour bien apprendre 5

SAVOIRS 7

Thème 1 - Structures algébriques usuelles 8

Thème 2 - Espaces vectoriels et applications linéaires 27

Thème 3 - Calcul matriciel 40

Thème 4 - Réduction des endomorphismes 59

Thème 5 - Espaces préhilbertiens réels 69

Thème 6 - Endomorphismes des espaces euclidiens 78

Thème 7 - Espaces vectoriels normés 84

Thème 8 - Suites numériques 98

Thème 9 - Fonctions réelles d'une variable réelle 102

Thème 10 - Dérivation et intégration des fonctions vectorielles 107

Thème 11 - Séries et familles sommables 119

Thème 12 - Suites et séries de fonctions 129

Thème 13 - Séries entières 135

Thème 14 - Intégration sur un intervalle quelconque 141

Thème 15 - Espaces probabilisés 151

Thème 16 - Variables aléatoires discrètes 158

Thème 17 - Équations différentielles linéaires 170

Thème 18 - Calcul différentiel 179

SAVOIR-FAIRE 187

Thème 1 - Structures algébriques usuelles 188

Thème 2 - Espaces vectoriels et applications linéaires 214

Thème 3 - Calcul matriciel 249

Thème 4 - Réduction des endomorphismes 288

Thème 5 - Espaces préhilbertiens 332

Thème 6 - Endomorphismes des espaces euclidiens 355

Thème 7 - Espaces vectoriels normés 369

Thème 8 - Suites numériques 398

Thème 9 - Fonctions réelles d'une variable réelle 418

Thème 10 - Dérivation et intégration des fonctions vectorielles . 430

Thème 11 - Séries et familles sommables 453

Thème 12 - Suites et séries de fonctions 479

Thème 13 - Séries entières 510

Thème 14 - Intégration sur un intervalle quelconque 536

Thème 15 - Espaces probabilisés 568

Thème 16 - Variables aléatoires discrètes 587

Thème 17 - Équations différentielles linéaires 613

Thème 18 - Calcul différentiel 637

CORRIGÉS DES EXERCICES 653

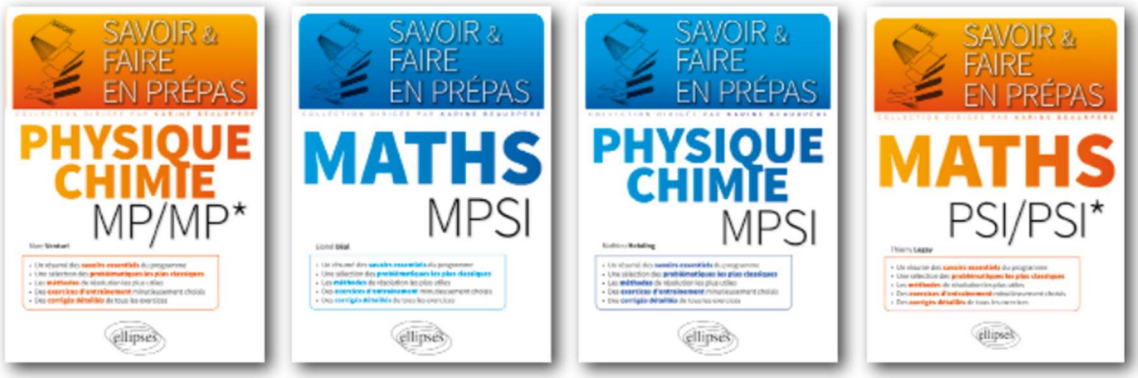
SAVOIR & FAIRE EN PRÉPAS

La collection ***Savoir et Faire en Prépas*** fait la promesse suivante aux étudiants des classes préparatoires aux grandes écoles (C.P.G.E.) : **cerner les savoirs primordiaux et acquérir les bonnes méthodes** pour faire face aux problématiques les plus fréquemment rencontrées.

Pour cela, la collection met en avant les savoirs essentiels mais elle met surtout en lumière la façon dont sont utilisés ces savoirs au travers de nombreuses méthodes.

Les **exercices corrigés** en fin de chapitres permettent de s'assurer d'avoir acquis la théorie et la pratique.

Cette collection démontre finalement que la résolution d'une grande quantité de problèmes se fait avec assez peu de savoirs théoriques mais beaucoup de savoir-faire.



9 782340 014831

www.editions-ellipses.fr